

Tier II intervention in the intermediate grades:
The effects of reciprocal teaching (RT) on standardized literacy assessments

by

Theodora D. Roop

B.S., Emporia State University, 2007
M.S., Emporia State University, 2010

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Curriculum and Instruction
College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Abstract

The purpose for this study is to look at the effects of a specific comprehension strategy—Reciprocal Teaching (RT; Palincsar & Brown, 1983,1984) on the performance of intermediate elementary students in the intermediate elementary grades on a standardized literacy assessment. The study was carried out in two implementation stages: a pilot with a small sample of fourth graders and a full-scale study with third, fourth, and fifth graders in a suburban public school district in the Midwest. The pilot study was implemented in two fourth grade classrooms, a total of 39 students, nine of which met the criterial for the experimental group who scored below the 25th percentile on an AIMSweb Maze 2016 fall benchmark. The full-scale study consisted of 269 students in grades three, four, and five across the same district, except for the fourth graders who previously were involved in the pilot. There were 71 students who met the criteria for the experimental group of scores below the 25th percentile on an AIMSweb Maze 2016 winter benchmark; however, after signed consent forms were returned, 59 participated. The students in the experimental group participated in intervention utilizing RT for approximately four weeks in Tier II small groups, instructed by classroom teachers at each elementary school.

Regression discontinuity was used to determine the effect of RT on the students' comprehension as indicated by their spring benchmark scores. The results of the pilot study showed a significant main effect in the scores along the regression line, indicating a positive effect for the intervention. Regression discontinuity analysis between groups indicated that a significant main effect exists for the intervention group. The full-scale study examined the effects per grade level for all students and with a set caliper for students whose scores were four points above and below the cut-off point, or winter 2016

benchmark, for the grade level. The results of the full-scale study focusing on all scores showed a significant main effect in the scores along the regression line, indicating a positive effect for the intervention for all grades; the effect was significant for all grades, but mostly for fourth and fifth graders. Regression discontinuity analysis between groups indicated that a significant main effect exists for the intervention group. When looking at scores within the caliper of four points above or below the benchmark, all scores showed a significant main effect in the scores along the regression line, indicating a positive effect for the intervention for all grades. Regression discontinuity analysis between groups indicated that a significant main effect exists for the intervention group within the caliper at each grade level.

Tier II intervention in the intermediate grades:
The effects of reciprocal teaching (RT) on standardized literacy assessments

by

Theodora D. Roop

B.S., Emporia State University, 2007
M.S., Emporia State University, 2010

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Curriculum and Instruction
College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. F. Todd Goodson

Copyright

© Theodora D. Roop 2018.

Abstract

The purpose for this study is to look at the effects of a specific comprehension strategy—Reciprocal Teaching (RT; Palincsar & Brown, 1983,1984) on the performance of intermediate elementary students in the intermediate elementary grades on a standardized literacy assessment. The study was carried out in two implementational stages: a pilot with a small sample of fourth graders and a full-scale study with third, fourth, and fifth graders in a suburban public school district in the Midwest. The pilot study was implemented in two fourth grade classrooms, a total of 39 students, nine of which met the criterion for the experimental group who scored below the 25th percentile on an AIMSweb Maze 2016 fall benchmark. The full-scale study consisted of 269 students in grades three, four, and five across the same district, except for the fourth graders who previously were involved in the pilot. There were 71 students who met the criteria for the experimental group of scores below the 25th percentile on an AIMSweb Maze 2016 winter benchmark; however, after signed consent forms were returned, 59 participated. The students in the experimental group participated in intervention utilizing RT for approximately four weeks in Tier II small groups, instructed by classroom teachers at each elementary school.

Regression discontinuity was used to determine the effect of RT on the students' comprehension as indicated by their spring benchmark scores. The results of the pilot study showed a significant main effect in the scores along the regression line, indicating a positive effect for the intervention. Regression discontinuity analysis between groups indicated that a significant main effect exists for the intervention group. The full-scale study examined the effects per grade level for all students and with a set caliper for students whose scores were four points above and below the cut-off point, or winter 2016

benchmark, for the grade level. The results of the full-scale study focusing on all scores showed a significant main effect in the scores along the regression line, indicating a positive effect for the intervention for all grades; the effect was significant for all grades, but mostly for fourth and fifth graders. Regression discontinuity analysis between groups indicated that a significant main effect exists for the intervention group. When looking at scores within the caliper of four points above or below the benchmark, all scores showed a significant main effect in the scores along the regression line, indicating a positive effect for the intervention for all grades. Regression discontinuity analysis between groups indicated that a significant main effect exists for the intervention group within the caliper at each grade level.

Table of Contents

List of Figures	xii
List of Tables	xiii
Acknowledgements	xiv
Dedication	xv
Chapter 1 - Introduction	1
Statement of the Problem	2
Purpose of the Study	3
Research Questions	5
Significance of the Study	6
Definition of Terms	7
Defining the Study	9
Limitations of the Study	9
Significance of Limitations	10
Organization of the Study	11
Chapter 2 - Literature Review	13
The Gist of RT	15
Direct instruction.	16
<i>Small group instruction.</i>	17
<i>Modeling.</i>	18
<i>Gradual release of responsibility and scaffolding.</i>	18
Developing the Metacognitive Mind of a Reader	21
Metacognition.	22
<i>Modeling metacognition via think-alouds.</i>	23
Predicting	24
Questioning and clarifying	25
Summarizing	26
RT and assessments.	27
Summary	28
Chapter 3 - Methodology	29

General Overview	29
Research Questions	29
Method	30
Participants and setting.	30
Participant selection procedure.	30
Materials and measurements	31
Experimental variable and general procedure.....	32
Descriptions of intervention.....	32
Research design.	33
Analysis.....	35
Validity.	37
Pilot Study.....	38
Method	38
<i>Participants and setting</i>	38
<i>Participant selection procedure</i>	38
<i>Materials and measurements.</i>	41
<i>Experimental variable and general procedure.</i>	41
<i>Research design.</i>	41
Full-Scale Study.....	42
Method	42
<i>Participants and setting.</i>	42
<i>Participant selection procedure</i>	42
<i>Materials and measurements.</i>	42
<i>Experimental variable and general procedure.</i>	43
<i>Research design.</i>	44
Summary	45
Chapter 4 - Results.....	47
The Sights and Sounds During Intervention.....	47
Overview of daily lesson plans.	47
Lesson sequencing and organization.	48
Classroom setting.....	50

Pilot Study.....	51
Full-Scale Study.....	53
Effects per grade level within the caliper.....	56
<i>Third grade.</i>	56
<i>Fourth grade.</i>	61
<i>Fifth grade.</i>	63
Observer agreement.	66
Summary.....	68
Chapter 5 - Conclusion	70
Discussion of Findings.....	70
Effects of RT as Targeted Instruction on Standardized Literacy Assessments	71
Differences in the Effect RT has on Standardized Test Scores at Each Grade	72
Increase in Overall AIMSweb Maze Scores due to RT as an Intervention	73
Importance to the Education Field.....	74
Limitations	75
Future Research	76
Classroom-based implications	76
Research-based implications.....	77
Conclusion	79
References.....	81
Appendix A - Sample AIMSweb Testing Passage	91
Appendix B - Introductory Lesson Plan	92
Appendix C - Basic Daily Lesson Format	96
Appendix D - Lesson Planning Guide	100
Appendix E - Observation Form.....	101
Appendix F - Interest Inventory.....	102
Appendix G - Teacher Feedback Chart	104
Appendix H - Non-Fiction Text Sample.....	107
Appendix I - Anchor Charts.....	109
Appendix J - Role Cards	110
Appendix K - RT Dry-Erase Mat.....	111

Appendix L - RT Spinner Game.....	112
Appendix M - Student/Teacher Graphic Organizer.....	113

List of Figures

Figure 1.1. CCSS English Language Arts Anchor Standards.....	6
Figure 4.1. Third Grade Discontinuation at Cut-Off Point Comparing Pre-Test Group Scores and Post-Test Group Scores Following Treatment.....	60
Figure 4.2. Fourth Grade Discontinuation at Cut-Off Point Comparing Pre-Test Group Scores and Post-Test Group Scores Following Treatment.	63
Figure 4.3. Fifth Grade Discontinuation at Cut-Off Point Comparing Pre-Test Group Scores and Post-Test Group Scores Following Treatment.....	65

List of Tables

Table 3.1 District Demographic Information, 2016-2017	30
Table 3.2 AIMSweb R-CBM and Maze Testing, Grade 4 Raw Scores	40
Table 3.3 Fourth-grade students participating in the RT intervention.....	40
Table 3.4 Full-Scale Study Student Demographic Information.....	42
Table 4.1 Pilot Study Student Demographic Information	51
Table 4.2 AIMSweb R-CBM and Maze Testing, Grade 4 Raw Scores	52
Table 4.3 Effects on Post-Test Scores	53
Table 4.4 Full-Scale Study Student Demographic Information.....	54
Table 4.5 AIMSweb Maze, Grades 3-5	55
Table 4.6 AIMSweb Maze, Grades 3-5	56
Table 4.7 Third Grade Gain Scores	59
Table 4.8 Fourth Grade Gain Scores.....	61
Table 4.9 Fifth Grade Gain Scores	64
Table 4.10 Effects on Post-Test Scores within Caliper per Grade Level	66
Table 4.11 Intervention Lesson Observations--Observer Agreement	67
Table 4.12 Teacher, Observer, and Administrator Exit Survey.....	68

Acknowledgements

I would not be able to be who I am and where I am in my life without my Lord, Jesus Christ, who strengthens me and with whom all things are possible. I have been blessed and privileged to be surrounded by some amazing and extraordinary people who have guided and supported me. With love and appreciation for all their sacrifices, to my first teachers--my parents, grandparents, and great-great parents, who instilled the love of learning and curiosity, and the will to persevere through challenges. I am grateful for the dedication of my teachers, from whom I learned that teaching is more than a job—it is the calling that can change the world, one child at a time.

Thank you to my husband, Mike, and my kids, Phoebe and Brodi, for your unending support and for the inspiration to keep me going with my “big homework project.” Words cannot express how much I love you all! The elementary teachers and administrators with whom I worked and learned so much from. My BCC and ESU colleagues, thank you for your continued encouragement; much appreciation to Dr. Clint Stephens for his guidance with the Regression Discontinuity design. Thank you to my friends and fellow “nerds” Kathy, Melissa, Amber, Jill, and Sara.

Thank you to my dissertation committee, my major professor Dr. F. Todd Goodson, Dr. Lori Goodson, Dr. Vicki Sherbert, Dr. John Schlup, and Dr. James Teagarden. I am truly grateful for your words of encouragement, guidance, and patience. My experience at K-State has been amazing! A big thank you to Dr. Mickey Losinski —where the idea for this study began to unfold and where I (almost) overcame my fear of quantitative research.

Dedication

This dissertation is dedicated to the teachers who empower students through literacy.

Chapter 1 - Introduction

As an elementary school teacher, I spent one year in second grade and six years in first. For me, first-graders will always be “fantastic”: anticipating the sacred time of recess, cherishing every moment of show-and-tell, playing with real and imaginary friends, while breaking the written language code, learning to read, and writing opinion pieces. They learn a lot in a year. They persevere through abstract concepts such as learning to read, and for many the process of learning to read is laborious and frustrating. Whether early exposure to reading, maturing level, or differences in rate of learning, students, not just my “fantastic” first-graders, struggle with different components of reading such as decoding, fluency, or comprehension. Before they can meet the grade-level standards, they have to step their way up through a deficit of subskills; they have to receive the appropriate intervention, not next month, not next semester, not next year, but right now.

My colleagues and I felt that we had the assessment tools to reveal individual student’s needs; we also had a plethora of resources in the classroom, in the reading lab, or online. In my mind there was an urgency of finding the right match between the need and the intervention resource but there was also the challenge of finding the time to investigate its effectiveness. I was only one person with a long to-do list, teaching every subject in a classroom of diverse learners. How can I be sure that what I implement for my most needy readers, will be the right intervention—one that has been put to the test? I have to equip them with the necessary building-blocks of learning to read, so that they can be readers who learn. After all, they will have to face the state-assessments, the complexity of content text in middle school and high school, and the expectations of society.

Since *A Nation at Risk* (U.S. Department of Education, 1983), the panic over the failure of public schools to adequately educate students, reached its culmination in the enactment of the 2001 No Child Left Behind (NCLB) mandates. Following this law, K-12 education in the United States has experienced the wake of strict accountability and achievement targets dominating the field for over a decade (Darling-Hammond, 2007; Bunch, 2011; Dee & Jacob, 2011; NCLB, 2001). The current trend of accountability measures found in the Common Core State Standards ([CCSS], National Governors Association [NGA], 2015/2017; Council of Chief State School Officers [CCSSO], 2010/2017) is geared towards a unified thread of what to teach rather than how to teach. While schools clearly know the expectations for the curriculum, benchmarks performance, and goals for achievement, the question of how to reach these is preset and urgent. Many commercial publishers advertise and label products and scripted programs as CCSS aligned. Teachers have also created a movement of internet based information exchanges via blogs, wikis, and websites where CCSS units and ideas are sold for a small profit. Nevertheless, each decision made about student progress and achievement must be cultivated upon data through continuous assessment. It should also be supported with research-based, systematic, and appropriate instruction, especially for students who are struggling, at-risk, and receiving targeted instruction as Response to Intervention ([RtI], National Center for Learning Disabilities [NCLD], 2017).

Statement of the Problem

Despite physical stamina, cognitive development, environmental factors, cultural diversity, and social-emotional characteristics, student academic performance is still a measure of accountability and an expectation for school excellence. While there are a variety of commercially produced resources purchased by school districts and teacher-created publications

that are easily accessible online, discerning research-based effective interventions for struggling readers can be a challenge. To achieve positive academic outcomes, human resources, time, and expertise in diagnosing and remediating reading difficulties are necessary to accurately assess student performance and to plan data-driven, appropriate instruction.

Students whose scores suggest below grade level performance based on standardized assessments, may show growth over a period; however, they are continuously racing towards a moving target of benchmarks and levels of learning mastery. Before students can achieve the expectations for mastery set forth by the CCSS ([NGA], 2015/2017; [CCSSO], 2010/2017), they need the necessary tools, strategies, and skills required for learning in a particular grade level. Addressing difficulties in learning, could mean examining the effectiveness of curriculum resources and instructional practices for comprehension of content area text that will reach the goal of desired academic performance.

Purpose of the Study

A walk-through in an elementary classroom can reveal the daily literacy experiences all students are exposed to: whole group reading with the teacher, small groups of students collaborating, or independent writing at a desk. What is obvious is that there is an established literacy curriculum but the ability of each child to meet the grade-level academic demands and expectations is not in plain sight. Assessments that measure the development of literacy skills to the ability of students to perform these successfully are used in many districts, if data is to drive instruction. Resources from adopted literacy curriculum are also in place. The disconnect between identified needs from the assessments and the availability of research-based tools for intervention can be an issue related to the lack of funding, trained reading specialist, or time to collaborate, plan, and prepare.

The purpose for this study is to look at the effects of a specific comprehension strategy on the performance of intermediate elementary students in grades three, four, and five on a standardized literacy assessment. Reciprocal Teaching (RT) was selected as the strategy for interventions. RT is a self-monitoring reading comprehension strategy that utilizes the components of summarizing, questioning, clarifying, and predicting (Palincsar & Brown, 1983,1984). It is peer-lead and teacher-guided approach to increasing student's self-monitoring of their comprehension before, during, and after reading.

Several studies report significant increases and statistically significant gains in scores on standardized tests for reading comprehension following RT implementation in the intermediate grades when RT was utilized (Fung, Wilkinson, & Moore, 2003; Oddo, Barnett, Haskins, & Musti-Rao, 2010; Sporer, Brunstein, & Kieschke, 2009). However, McKewon, Beck, and Ronette (2009) report no significant effect of strategy use, such as RT, and recommend that the focus is shifted from strategy instruction to understanding of how concepts such as summarizing and predicting aid overall reading comprehension, even before students can decode. How are such discrepancies, often contracting statements or conclusions, addressed when making decisions about targeted instruction?

The study further explores previous research by:

1. Utilizing Palincsar and Brown's (1983, 1984) RT study as an instructional tool for targeted instruction during small group, Tier II intervention. In previous studies, RT has been mainly used with whole or small groups of students. This study revisits RT as an instructional resource for reading intervention based on current assessment tools, with different intermediate grade levels participating in the intervention simultaneously under a Regression Discontinuity (RD), quasi-experimental design.

2. Bridging reading curriculum resources. These can vary from one school district to another, and from state to state; however, the CCSS ([NGA], 2015/2017; [CCSSO], 2010/2017) are adopted by 42 states. The selection of non-fiction texts used in the study are not from reading series, basal readers, or programs; however, they are aligned with the current standards, which are listed with the text selections, and serve as the medium for strategic reading for learning in the content areas.

3. Integrating learning strategies/skills and background knowledge between Tier I and Tier II. While providing a systematic lesson plan protocol for RT's strategies, students and teachers can refer to and integrate previously mastered, Tier I curriculum methods for decoding, forming questions, locating text features, or finding main ideas as they predict, question, clarify, and summarize.

Research Questions

Making decisions about instructional resources that bridge the gap between deficiency in reading-to-learn and mastery of standards for reading comprehension are critical for school administrators, interventionists, and classroom teachers. Immediate intervention and limited resources demand the selection of appropriate intervention instruction that spurs growth in reading scores in the most efficient and effective manner. The following research questions are designed to measure the effectiveness of utilizing targeted instruction, such as RT, for students in grades three, four, and five.

1. What are the effects of Reciprocal Teaching (RT), as targeted instruction, on standardized literacy assessments for students in the intermediate grades?
2. Are there differences in the effect, if any, RT has on students' standardized test scores at each grade level?

3. Does RT increase students' overall scores on AIMSweb Maze over a period of targeted intervention?

Significance of the Study

The significance of this study strives to address several challenges schools face today linked to academic achievement such as adopting effective and efficient instruction and limited research-based resources for early intervention. Students who struggle to comprehend in the intermediate grades are reading to learn rather than learning to read. They are transitioning to more non-fiction and content area reading.

Current CCSS requirements intend for students to be active and engaged readers attending to in-depth analysis of wide range of texts. Students are expected to independently and proficiently determine main ideas, to analyze, and to evaluate texts, which are cognitive tasks beyond basic comprehension (see Figure 1.1). When students are identified below grade level, based on standardized tests, the intervention they are exposed to must be effective and efficient to address these academic needs. Schools must make informed decision about successful interventions, as immediate students' needs must be addressed with available resources that will provide them with opportunities to reach CCSS mastery.

Figure 1.1. CCSS English Language Arts Anchor Standards

Key Ideas and Details	Craft and Structure
CCSS.ELA-LITERACY.CCRA.R.1 Read closely to determine what the text says explicitly and to make logical inferences from it; cite specific textual evidence when writing or speaking to support conclusions drawn from the text.	CCSS.ELA-LITERACY.CCRA.R.4 Interpret words and phrases as they are used in a text, including determining technical, connotative, and figurative meanings, and analyze how specific word choices shape meaning or tone.
CCSS.ELA-LITERACY.CCRA.R.2 Determine central ideas or themes of a text and analyze their development; summarize the key supporting details and ideas.	CCSS.ELA-LITERACY.CCRA.R.5 Analyze the structure of texts, including how specific sentences, paragraphs, and larger portions of the text (e.g., a section, chapter, scene, or stanza) relate to each other and the whole.
CCSS.ELA-LITERACY.CCRA.R.3 Analyze how and why individuals, events, or	

ideas develop and interact over the course of a text.	CCSS.ELA-LITERACY.CCRA.R.6 Assess how point of view or purpose shapes the content and style of a text.
Integration of Knowledge and Ideas	Range of Reading Level of Text Complexity
CCSS.ELA-LITERACY.CCRA.R.7 Integrate and evaluate content presented in diverse media and formats, including visually and quantitatively, as well as in words. ¹ CCSS.ELA-LITERACY.CCRA.R.8 Delineate and evaluate the argument and specific claims in a text, including the validity of the reasoning as well as the relevance and sufficiency of the evidence. CCSS.ELA-LITERACY.CCRA.R.9 Analyze how two or more texts address similar themes or topics in order to build knowledge or to compare the approaches the authors take.	CCSS.ELA-LITERACY.CCRA.R.10 Read and comprehend complex literary and informational texts independently and proficiently.

Common Core State Standards Initiative (CCSSI, 2017)

Definition of Terms

1. AIMSweb Maze: standardized assessment developed by Pearson Education, Inc¹. (2012); utilized by school districts as a screening and progress monitoring assessment tool. This component focuses on measuring students' ability to monitor their reading comprehension. After the first sentence, every seventh word is omitted and the students must choose the best word from three given words to go in the blank. The students have three minutes to complete the cloze-passage; only correct responses are counted toward the score.
2. Cut-Off Score: also, referred to as precut but is different from the grade level benchmark; this is the predetermined score by grade level, either by Pearson Education, Inc., the district, or both. This score varies by grade level.

¹ The study was funded entirely by the researcher; Pearson Education Inc. did not contribute to this study monetarily or in any other form.

3. Dependent Variables (DV)-- Posttest Scores: based on AIMSweb Maze benchmark after the intervention.
4. Effect: refers to the statistical significance of an outcome, calculated as a p-value or the probability of observing an effect given that the null hypothesis is true. For a $p < 0.05$, the probability of observing an effect given that the null is true is less than 5%.
5. Independent Variables (IV)--Pretest scores: based on AIMSweb Maze benchmark prior to intervention.
6. Intermediate Grades: grades at the elementary school level, specifically third, fourth, and fifth grade. Students approximate age range is 8-9 for third, 9-10 for fourth, and 10-11 years old for fifth grade.
7. Reciprocal Teaching (RT): a self-monitoring reading comprehension strategy that utilizes the components of summarizing, questioning, clarifying, and predicting (Palincsar & Brown, 1983,1984).
8. Standardized Literacy Assessments: assessments or tests with the same protocol, directions, questions/requirements for large groups of test-takers; scored in a consistent fashion for comparison across individuals or groups of individuals.
9. Standardized Scores: raw scores from a standardized assessment.
10. Targeted Instruction: based on data from assessments, this is an intentional, needs-based lesson protocol, aiming to address specific skills, such as reading comprehension.
11. Targeted Intervention: refers to the designated instructional time of consistent and systematic delivery of targeted instruction.

Defining the Study

The present study is quantitative in nature. It explores a given problem, as the one defined above, in an experimental setting and by defining a hypothesis about the outcome following observation, data collection, and data analysis. It utilizes a quasi-experimental design to investigate the effectiveness of the treatment applied in the experimental group, with data from pre- and post-intervention testing. The design does not employ a random assignment in its true form; it however, mimics it when comparing students in the experimental and in the control groups.

The study is split into a pilot and a full-scale stage. Both were conducted in elementary school buildings with students in grades three, four, and five. The purpose for the pilot study was exploratory and only with fourth grade participants. The need for several adjustments for the full-scale were discovered in the process of conducting the pilot study, such as to the design, analysis, and materials and resources management. The full-scale study included 59 elementary students in grades three, four², and five, who were given an interest inventory for the purpose of motivation. The analysis included comparing the experimental and control groups with all students in those grades and with students within a caliper, who are assumed to be similar in academic performance on the experimental variable. The gathering, distribution, and materials management format from the pilot study worked efficiently for the full-scale study.

Limitations of the Study

While the general purpose of the study appeals to contemporary trends in education, it also presents several limitations. The pilot and the full-scale study were conducted in a Midwest, suburban area, where much of the population is Caucasian. Little diversity in race and ethnicity

² The fourth-grade students and teachers from the pilot study did not participate in the full-scale study.

exists, although 53% of the population in the community earns an annual income below the median of \$58,000³.

The researcher of the study is a former classroom teacher with the perspective that the conclusions of the study are meant for empowering students learning by equipping elementary classroom teachers with effective intervention instruction. Differences in pedagogical philosophy, district vision and mission, or specific academic and demographic needs may not align with the recommendations for implementation or scope of significance presented.

The study is also conducted with the participation of minors, which requires a consent form signed by a parent or a guardian. It is possible and realistic to expect that in some cases, parent or guardian consent to participate in the intervention may not be granted for reasons other than those within the control of the study. For the purpose of broad generalizations, representative of the entire intermediate grades school population in the US, results should be interpreted with caution in the case of a smaller than expected number of consents are granted.

Significance of Limitations

The time of the year, before and after a winter break and before summer break, can also pose several obstacles to the scheduling and the study. The teachers are preoccupied with the last few days of testing and end-of-school events; the students are anticipating the break from school and enjoying the spring weather which can be distracting for engagement with academic tasks.

Researcher bias and teacher perspectives can both influence the interpretation of the effect and the student ability for summarizing can be subjective based on students' background knowledge and perspective on the topic. It can also be influence by the teacher and what they

³ Values are rounded estimates to avoid identifying information, preserve confidentiality, and are based on 2015 data.

find to be important in the text when summarizing volumes of information from text. Teacher experience with strategy instruction will also vary between years of teaching and years of teaching a specific grade level.

Students who meet the criteria but do not return signed consent forms (or permission was not granted), even after teacher follow up in person or via written communication can lead to decreased number of students participating in the experiment group in the full-scale study. There may not be enough statistical power to detect significance for the particular quasi-experimental design with low number of participants.

Organization of the Study

At present, Chapter 1 displayed the purpose of the study and questions aimed at determining the impact of RT as an instructional resource for Tier II intervention. Several reasons for the significance of the study as well as the scope of limitations were posed. A socio-cultural perspective is threaded throughout the review of literature in Chapter 2. It contains previous studies and literature examining RT, including making predictions, formulating questions, clarifying at the linguistic and concept-oriented level, and summarizing. Metacognition and instructional delivery, along with its underlying concepts are also examined in relation to RT.

The methodology for the pilot and full-scale stages of study is discussed in Chapter 3. The participants, selection, and setting are described. This chapter proceeds to explain the design and analysis conducted in both stages. A review of the outcome variable, validity, and statistical model follows.

Chapter 4 presents the results of the pilot and the full-scale stages of the study. This section organizes the empirical data gathered over both stages of the study. Statistical

calculations are organized and analyzed for practical interpretation. Each research question is addressed considering the results of the study.

The consequent Chapter five, articulates on the conclusions and suggests several implications for the classroom and for future research. The limitations and the importance of the study for the field of education are revisited by reflecting on the results.

Chapter 2 - Literature Review

Literacy has revolutionized human consciousness for centuries. The shift from orality based cultures to mass literacy in contemporary cultures has created an ongoing analysis in the twentieth century, particularly in the early 1980's (Ong, 1982; Pattison, 1982; Havelock, 1986). During the late 1970s and early 1980's the whole language movement encountered *A Nation at Risk* (U.S. Department of Education, 1983)—a reflection of the failing educational system calling for education reforms with Reagan's "back to basics." The dilemma of what the perfect reading curriculum should teach the ever-changing and diverse class-scape in the schools has been a subject of interest for policy makers, researchers, and educators. The National Reading Panel ([NRP], National Institute for Child and Human Development, [NICHD], 2000) identified five core areas for reading instruction: phonemic awareness, phonics, fluency, vocabulary, and comprehension. Each year, students are subjected to multiple assessments, directly or indirectly addressing each of these areas; these are meant to monitor student progress and assure success on end-of-the year high-stakes testing. This chapter examines literature of past and current trends in reading instruction for comprehension. The focus is on the strategy of Reciprocal Teaching (RT), which consists of predicting, questioning, clarifying, and summarizing with text. Each component is examined in terms of significance to increasing comprehension and overall metacognition. The method of direct instruction is dissected into modeling, scaffolding, and gradual release of responsibility to review the effectiveness of these delivery tools when using RT for comprehension instruction.

Keene and Zimmermann (2007) differentiate between surface structures and deep structures of the text. Being able to read the words of the text, that are written with letters and in a specific grammatical pattern, is not enough to gain understanding of the text. Understanding

the text alone is only at the basic level of comprehension. Tovani's (2000) conversation with Luke, a ninth grader and a struggling reader, is a reminder of why effective and timely instruction of those deeper structures of comprehension is necessary. When Luke struggled to understand what he had read, he explained that, "When I was younger, I used to try sounding words out but that didn't help [or] I don't do anything" (p. 49). The extent of his strategic reading when meaning breaks down is alarming as he continued to reflect, "Nope, I keep reading and hope it makes sense when I am done [and when it doesn't make sense], then, oh well" (p. 49). Before students are immersed in the vast amounts of content reading in the middle and high school grades, they must be engaged and cognizant readers; they should not only understand what they read but know when they lack understanding. Early intervention with effective instructional tools and strategies was needed for students like Luke to adopt the mindset of a reader, who can recognize the deeper structures of the text and expand their learning capacity.

The goal of RT is to provide the reader with a deeper understanding and allow for critical thinking, and to further curiosity and learning growth from text; it is "a tool not an end in itself" (Dewitz & Graves, 2014, p. 156). Strategies themselves, and in this case RT, should be the means not the end of comprehension instruction (Jacob & Paris, 1987; McKeown et al., 2009). Rooted in the Vygotskian socio-cultural perspective, RT initiates learning to occur from teacher modeling and guidance through peer interaction, leading to internalization of the skill needed for independent performance. Vygotsky (1978) proposed that,

When it was first shown that the capability of children with equal levels of mental development to learn under a teacher's guidance varied to a high degree, it became apparent that those children were not mentally the same age and that the subsequent course of their learning would obviously be different. The difference between twelve and

eight, or between nine and eight, is what we call the zone of proximal development. It is the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers. (p. 79)

The social learning aspect of peer teaching, mimics the language discourse taught by the expert reader, or the teacher, through modeling and gradual release, with opportunities for practice in a small group setting, with more capable peers within each child's zone of proximal development (ZPD). The students are not only engaged readers but active listeners, critical thinkers, and problem solvers with a common goal.

The Gist of RT

Reciprocal Teaching (RT) was first introduced in a report of the seminal work by Palincsar and Brown (1983, 1984). The three studies described in the report, were conducted with 7th grade students, who were primarily good decoders but poor comprehenders. The strategies of RT as presented in the report were questioning, predicting, summarizing, and clarifying (Palincsar & Brown, 1983,1984). Following a direct instruction format the teacher first introduced and modeled the strategies one at time (Oczkus, 2010; Alvermann, Gillis, & Phelps, 2013). With the teacher's guidance and feedback, instruction moved from "delivery to discovery" for the students (Fisher, Brozo, Frey, & Ivey, 2007, pp. 89-92). Students have the opportunity to mimic the natural occurring dialogue with appropriate scaffolding and gradual release of responsibility (Pearson & Gallagher, 1983; Alvermann, et al., 2013).

In the classroom, RT can be used in whole group, small groups, or during literacy circles (Oczkus, 2010). The instructional cycle begins with teacher modeling of the strategies one at a time, thinking out loud about how the strategy helps the reader connect with the text. Graphic

organizers can be used to record the information, with evidence from the text, supporting the reasoning for predicting, questioning, clarifying, or summarizing. Depending on the length of the text and age of the students, the teacher can focus on a paragraph, a section, or an entire chapter at a time.

The students are to mimic the teacher during guided practice, when the role of the teacher is to gradually allow the students to utilize the strategies more and to provide feedback and support when needed. During this guided practice, the teacher monitors the class or group, to assess who is ready for more responsibility. Each student takes on a different role of predictor, questioner, clarifier, or summarizer; this can be done in groups of four if working with the whole class or within the small group of 4-6 students.

Before students engage in independent practice, they spend time working with a peer. The reciprocal part of RT occurs at this stage, when pairs of students, still under some teacher supervision, practice the same modeled behaviors and metacognitive language, with text that they can read fluently, so that all their cognitive energy is channeled to comprehension. The learning from text happens in a social setting, with the students taking on the roles of predictor, questioner, clarifier, or summarizer. The students make decisions about which strategy is used and when to use it appropriately. They are collaborative as they work on the goal of learning from the text and they monitor their understanding while reading.

Direct instruction. Interestingly, direct instruction began its inception in advertising on studies examining input and retention. It was first used by Engelmann, as an instructional method with his children, later leading to research in early childhood settings and the development of programs in the 1960s in collaboration with Bereiter (National Institute for Direct Instruction [NIFDI], 2015). Rinehart, Stahl, and Erickson (1986) state that direct instruction is key to

training students in the metacognitive reading behaviors of self-regulation and monitoring. This can be effectively done in a short amount of time or via a short orientation (Raphael & McKinney, 1983; Rinehart et al., 1986). Meuller (2001) states that RT specifically, is a comprehension strategy that allows for meaning monitoring. It should be taught utilizing direct instruction and modeling by the teacher to the students in small groups of three to four, with texts at their independent or instructional level (Fisher, Frey & Lapp, 2009; Meuller, 2001).

Small group instruction. Small group instruction requires careful preparation and planning along with routines and procedures for teacher and peer interactions. Uttero (1988) provides the following steps for small group implementation: (1) state the guidelines; (2) organize groups of three to six students; (3) give an overview of the model to students; (4) model the strategies to the students; (5) utilize gradually release of responsibility; and (6) observe the interactions between students (pp. 394-395). Other studies by VanKeer and Verhaeghe (2005) with second and fifth graders, and Williams (2001) with suburban middle school students also utilized this small group lesson protocol along with scaffold instruction by the teachers for systematic utilization of simultaneous strategies. The texts were typically expository; however, RT can be successfully and effectively used with narrative and with variety of content areas and grouping formats, (i.e. whole class, small group/guided reading, one-on-one; Flippo, 2003; Elish-Piper & L'Allier, 2013).

The instruction during small groups should be explicit and systematic, especially when working with students who are at-risk and with ELL students (Baumann, 1984; Carnine, Kameenui, & Woolfson, 1982; Elish-Piper & L'Allier, 2013; Lipson & Wixson, 2009). Rosenshine and Meister (1994) found that explicit instruction on the RT strategies, prior to use of the strategies produced more significant results for students who were at-risk. Through direct,

systematic, and often scripted instruction, including teacher-guided application and independent practice, the strategy lessons should include introduction of the strategies, their purpose, and application examples (Baumann, 1986; Carnine et al., 1982).

Modeling. Modeling is an effective and crucial component of direct, explicit instruction (Baumann, 1984, 1986; Bean & Steenwyk, 1984). Specifically, modeling involves mimicry--the student watching the teacher as they effectively perform [and think aloud on] the use and the protocol of the strategy (Fisher, Frey, & Williams, 2002). Explicit instruction is key to success and growth with comprehension for students who are struggling and are placed in intervention groups (Brown, Campione, & Barclay, 1979). Teacher modeling of strategies should explicitly describe the strategy and when and how it is used (Tovani, 2000; Duke, Pearson, Strachan, & Billman, 2011). Intentional strategy instruction produced reading behaviors in students that “instead of reading nonstop through the passages, the students were overtly using strategies that increased their attention to the material. Students, to a varying degree, attended [to] subheadings, stopped to re-read, [and] reviewed after reading” (Adams, Carnine, and Gersten, 1982, pp. 49-50).

Gradual release of responsibility and scaffolding. Gradual release of responsibility ([GRR], Pearson & Gallagher, 1983) is another important component of direct instruction. It represents a varied amount of support and gives the teacher opportunities to make decisions about how much release of strategy application is given to the students; it also allows the teacher to provide necessary support to students through scaffolding (Davey, 1983). Part of the GRR includes teachers observing for automaticity of the skills taught, providing practice in various contexts, and striving to have similar context during instruction and during application

(Baumann, 1984; Dewitz & Graves, 2014). Duke et al. (2011) recommend that a gradual release of responsibility (GRR) is used when teaching any comprehension strategy.

Fisher and Frey (2014) describe the utilization of four types of scaffolding during the reading: “repeated readings, collaborative conversations, annotations, and text dependent questions” (p. 279). For struggling readers, they call for “contingency plans [consisting of] re-establishing purpose, analyzing questions to identify likely answer location, prompting and cueing, modeling [fix-up strategies], and analyzing annotations” (p. 282).

The amount of scaffolding can vary. As with other studies (Brown, Smiley, Day, & Lawton, 1977; Brown, Day, & Jones, 1983; Markman, 1977), Raphael and McKinney (1983) found that more prompting was required for younger students, while frequent prompting was more of a distraction for older students. This signals differentiated instructional consideration depending on the grade levels and the amount of skill proficiency. Finding implicit information required much more time of training than finding explicit information, which calls for shorter orientation, regardless of the age. In another study, Raphael and Wonnacott (1985), found that instruction duration on reading comprehension should vary based on the grade level—longer for younger students, and shorter for older students, but they caution that ability levels may also play a role. Raphael and Pearson (1985) found that average and low ability readers also required more prompting and support with prior knowledge, especially with finding the answers to implicit questions. Increased scaffolding by the teacher during RT instruction allowed students to stay focused on “their discussions of the text longer and more seriously, [they] were more active in their reading, and were provided with good models for summarizing, clarifying, predicting, and questioning” (Hacker, 2002, p. 703).

Another form of scaffolding can also be achieved by a mediator along the gradual release continuum. Graphic organizers are visual representations of the material and have been found to be an effective tool for strategy use in a group setting (Darsch, 1986; Fisher et al., 2002). Visual representations are “scaffolds provid[ing] us with both a technical vocabulary and tools for improving instruction” (Rosenshine, Meister, & Chapman, 1996, p. 211)

Vygotsky (1978), as previously quoted, emphasizes the level of support children need through “adult guidance or in collaboration with more capable peers” (p. 79). The level of expertise in planning appears to be the best predictor of efficiency as older students are more able to stay on topic and utilize prior knowledge (Brown et al., 1977; Brown et al., 1983); older students are also more aware of comprehension failure than younger students who tend to require prompting (Markman, 1977). At a latter level of the gradual release sequence, peer-teaching can also be viewed as what Vygotsky (1978) called “more capable peers” (p.79). Klingner, Vaugh, and Schumm’s (1998) study on cooperative learning with 4th graders, revealed that students were better able to apply procedural and conditional⁴ knowledge and had a better understanding of how to help each other.

Verbal and written expression of what students are comprehending through the strategy is part of the GRR during the guided practice component of direct instruction, as more responsibility is released from the teacher to the students, and they are "transforming information externally " (Fisher et al., 2009, p. 14). Oral language use, conversation, or the dialogue used in RT are mediums for the shared use of the strategy between the teacher and the student, with teacher support and guidance (Fisher et al., 2009). Such social interaction serves as the conductor

⁴ Procedural and conditional knowledge of strategy use refers to the awareness of how to use the strategy and when to use it effectively (see Jacob & Paris, 1987).

between language and thought; it is Vygotsky's (1985) counterargument to Piaget's theory of how language develops, and that is, language begins with social interaction rather than inner speech. He further explains that "our schema of development [is] first social, then egocentric, then inner speech" (p.19). It is only after much "practice" with language in a social setting that language is internalized and utilized for metacognition. The positive outcomes of the processes that occur through dialogue during RT can be seen in increased comprehension on many forms of assessments including standardized assessments (Lipson and Wixson, 2009; Vacca & Vacca, 2008). RT has been studied mostly within the intermediate and middle school levels (Palincsar & Brown, 1983, 1984); however, it can also be used in the primary classrooms with read-alouds due to the developing decoding skills for this group of students (Flippo, 2003; Lipson & Wixson, 2009). The lesson plan elements of modeling and guided practice with GRR and scaffolding, can be easily applied to RT. They are essential for the development of cognitive processing, by making sense of print through oral language and by receiving appropriate, immediate, and based on need feedback (Alvermann, et al., 2013; Flippo, 2003; Lipson & Wixson, 2009).

Developing the Metacognitive Mind of a Reader

It is important to differentiate between the strategies in RT as those, which would provide organization and initial comprehension, and those that help with monitoring once comprehension exists (Berkowitz, 1986). The RT strategies allow students to set a purpose for reading, activate background knowledge, formulate questions that pique their curiosity of the unknown, monitor their comprehension while reading, draw conclusions about parts of the text or for the whole composition, and make decisions about which strategy to use and when to use it to support their comprehension (Elish-Piper & L'Allier, 2013; Fisher & Frey, 2015; Flippo, 2003; Lipson & Wixson, 2009; Vacca & Vacca, 2008). Lipson and Wixson (2009) further explain that RT is "a

simulation of the reading comprehension process" (p. 556); the strategy is not the end but the means, by which students work on becoming strategic readers in small groups, as in Readers Workshop (Meuller, 2001), literature circles (Oczkus, 2010), or independently, across curriculum and for different purposes (Meuller, 2001). RT is a meaningful and authentic activity that has a real-life application of reading connected text and applying the strategies for learning in a social setting (Alvermann et al., 2013; Fisher & Frey, 2015).

Metacognition. Flavell (1979) refers to metacognition as the “actions and interactions” between four cognitive domains: “(a) metacognitive knowledge, (b) metacognitive experiences, (c) goals (or tasks), and (d) actions (or strategies)” (p. 906). Metacognitive knowledge can be conscious, and can be added to, removed from, or revised depending on metacognitive experiences. Based on these two domains, metacognitive experiences and metacognitive knowledge, one decides how to act (what strategy to use) based on their purpose (goal or task). These actions can be cognitive or metacognitive but the main point Flavell (1979) makes is that while understanding a message is important, evaluating the content of the message is critical for decision making. Being a metacognitive reader allows for monitoring one’s level of comprehension and deciding on effective strategy application. The reader becomes engaged in the process of learning by developing this self-awareness in order to critique the message rather than have only basic understanding (Mier, 1984). Zimmerman and O’Keefe (2007) describe the intentional use and awareness of strategic reading as the “framework” for not only understanding and content knowledge but more importantly for the ability of critical thinking. This view is also supported by McKeown, Beck, and Blake (2009) who suggest that modeling strategies within context, with short text, and having enough automaticity with declarative, procedural, and

cognition knowledge is necessary to derive the ideas and to show how they are connected in content reading.

Metacognition allows the reader to monitor and adjust one's own strategic engagement with text (Jacob and Paris, 1987). Harvey & Goudvis (2007) explain that instruction in metacognition should transform "tacit readers to aware, strategic, and reflective readers" (p. 26); furthermore, this instruction should be systematic (Flavell, 1979). Regarding metacognition, Jacob and Paris (1987) differentiate between two types of cognition: "self-appraisal of cognition" and "self-management of thinking" (p. 258). The former consists of the declarative, procedural and conditional knowledge of the strategic reader, and the latter deals with planning, evaluation, and regulation—in other words, monitoring. Klingner et al. (1998) place emphasis on teaching the procedural and conditional knowledge for strategies and assessing if students are using them effectively.

Modeling metacognition via think-alouds. Davey (1983) describes think alouds, or thoughts of a strategic reader, as verbalized statements by the teacher, who prior to reading aloud while students follow and listen, has identified possible areas "of difficulties, contradictions, ambiguities, or unknown words" (p. 45). Davey, (1983) suggest that when it comes to poor comprehension,

students often learn well when they can watch a good model [who] by verbalizing their own thoughts as they read-aloud model[s] the kinds of strategies a skilled reader uses during reading and pointing out specifically what they are doing to cope with a comprehension problem. (p. 44)

Prior to modeling, the teacher must be cognizant of their own reading behaviors to become aware of the processes involved that otherwise occur automatically for a strategic reader

(Smith, 1988). Before group practice, the teacher should model the metacognitive language for declarative, procedural, and cognitive knowledge described by Jacob & Paris (1987) based on the task and selection of strategies, as discussed by Flavell (1981). Jacob and Paris (1987) describe declarative knowledge as knowing what the strategy is, procedural as how to apply the strategy, and conditional as when and why to apply a specific strategy over another to be an effective strategic reader. Harvey & Goudvis (2007) refer to it as "shar[ing] our inner conversation [about] how we activate our background knowledge, ask questions, [and] draw conclusions" (p. 46). When doing so, the teacher should use "I" statements to demonstrate the negotiation of personal meaning as well as the purpose of intentional strategy use (Fisher and Frey, 2015; Jacob & Paris; 1987; Kilgner et al., 1998). In other words, modeling metacognition via think-alouds with text, provides students with the thinking process for self-monitoring and decisions about fix up strategies for meaning making (Davey, 1983; Zimmerman & O'Keefe, 2007). An essential part of direct instruction with RT should include the think-aloud modeling by teachers for each of the four strategies for fostering and self-monitoring of comprehension: predicting, questioning, clarifying, and summarizing.

Predicting. Previewing the text is an important step in comprehension instruction (Fisher & Frey, 2015). As part of this process, predicting sets a purpose for reading and creates awareness of thoughts when attempting to understand or comprehend; readers draw from schema, or prior knowledge, and connect it to what was read to understand or comprehend (Omanson, Beck, Voss, & McKeown, 1984; Raphael & Wonnacott, 1985). Students can draw from their "schemata [to] provide the interpretive framework for comprehending discourse" (Brown et al., 1977, abstract). Making predictions activates prior knowledge and connects it to what was read, allowing for comprehension monitoring (Harvey & Goudvis, 2007). Tovani

(2000) states that making predictions allows the reader to "anticipate" what comes next and to support comprehension during reading. When meaning is breaking down, the reader can revisit and revise the predictions and check for meaning by developing a new prediction (Tovani, 2000).

Questioning and clarifying. Questions can often arise from lack of clarity, derived from the awareness that comprehension is breaking down, or developed to search for specific information (Tovani, 2000; Harvey and Goudvis, 2007; Keene & Zimmermann, 2007). There are three distinct types of questions, depending on the level of comprehension or purpose for reading: literal, interpretive, and applied (Raygor & Raygor, 1985; Vacca & Vacca, 2008). Literal questions summon answers that come directly from the text, interpretive questions refer to making inferences, and applied questions promote forming opinions, new ideas, or thoughts.

The interdependence of question generation and clarification is related to the purpose of reading, especially for reading to learn. Readers should be taught to ask questions before, during, and after reading (Nolte & Singer, 1985; Uttero, 1988; Klingner et al., 1998; Keene & Zimmermann, 2007). Instruction in question developing should stress the importance of the question-answer relationship (Vacca & Vacca, 2008). If students are not aware of their own comprehension they are less likely to ask questions relative to the content and to search for clarification (Helfeldt & Lalik, 1976; Markman, 1977; Nolte & Singer, 1985). Questions prompt clarifications within a social learning setting, motivating students to find the answers (Nolte & Singer, 1985). Particularly with vocabulary, clarification may play a role in being able to delete trivial information for identifying main ideas (Bean & Steenwyk, 1984). Some forms of clarification with vocabulary may include looking at word parts, context and picture clues, and other fix-up strategies (Fisher & Frey, 2015).

With question generation, students are also more likely to learn what they ask about (Wixson, 1983). In Davey and McBride's (1986) study, 6th grade students who came up with their own questions, outperformed those who had to answer pre-formulated questions. In another study by Helfeldt and Lalik (1976), when students imitated the modeled behavior of teachers asking questions, they experienced more success responding to what they had read.

Summarizing. Readers can determine the important information, distinguish it from unimportant information, and commit it to long-term memory for further synthesis based on questions and purpose for reading, in a manageable and meaningful manner (Harris & Sipay, 1980; Keene & Zimmermann, 2007). Teaching students to use text structure and text features can help in locating important information based on the text rather than from schemata (Carnine et al., 1982). In Rinehart et al. (1986) study with sixth graders, students were better able to identify explicitly stated, within the paragraph main ideas, but had difficulty with those that were implicitly suggested in the text.

Main ideas in text are considered important information, which is tied to summarization (Baumann, 1984). The idea of using text structure and features, especially with expository text, is necessary for discovering implicit main ideas (see Rinehart et al., 1986). Instruction in text structure for information recall of expository text was applied successfully with 6th graders by Berkowitz (1986) when teaching study skills. For lower ability readers, the instruction should include much more scaffolding in the form of feedback and practice utilizing text structure to locate important information within expository and content area dependent text (Armbruster, Anderson, & Ostertag, 1987; Fisher & Frey, 2015). This may include the locating of headings and subheadings, as well as topic sentences (Bean & Steenwyk, 1984).

Each one of the strategies comprising RT contribute to the skill and effectiveness readers acquire to approach and support higher levels of comprehension from text. The level of success when reading to learn is measured on continuous basis over the course of instruction through teacher informal assessment. The culminating evaluation, however, is reflected on high-stakes, standardized test, designed to determine the level of academic achievement in the intermediate grades.

RT and assessments. Several studies report significant increases and statistically significant gains in scores on standardized tests for reading comprehension in the intermediate grades when RT was applied during instruction (Fung, Wilkinson, & Moore, 2003; Oddo et al., 2010; Lipson & Wixson, 2009; Sporer, Brunstein, & Kieschke, 2009). Even when no significant effect existed for comprehension scores following strategy instruction, it is recommended to explain to students how concepts such as RT's summarizing and predicting aid in overall reading comprehension (McKewon, Beck, and Ronette, 2009)

A major difference in the effectiveness of reciprocal teaching on standardized scores was found by Rosenshine & Meister (1994). The review consisted of nineteen studies on the strategy and its effect on student scores for different types of assessment. The overall conclusions on RT studies were that implicit instruction was not as effective when compared to studies that utilized explicit instruction prior to strategy utilization for comprehension. Teaching each component strategy of RT was essential as well as the gradual release from modeling, to guided practice, to peer interaction and dialogue. The main feature of RT was the dialogue that the teacher, when explicitly modeling, thinks out loud and allows the students to create a visual and spoken bank of vocabulary necessary for metacognitive processes to occur during guided practice and peer interaction. It was found that most often, students who were considered at risk (poor

comprehenders) received the implicit, while the rest received the explicit instruction. As for assessment, the results indicated more successful completion on researcher created assessments rather than standardized assessments. Rosenshine and Meister (1994) point out that the at-risk students were at a disadvantage--they were not only given more implicit instruction with RT but their progress was also measured on standardized assessments. There were no significant findings recommending the number of students in a group or duration of intervention.

Summary

Thus far, RT can be understood better as a process rather than a one-size-fits-all program or a set of isolated strategies. It involves the intentional and metacognitive engagement of the reader with the text before, during, and after the act of fluent decoding. Under direct instruction, teachers can serve as the model for metacognition with the text, while providing meaningful opportunities for practice and discussion. Learning occurs within a small group setting, allowing the teacher to closely monitor students' progress and provide appropriate scaffold to individual students. Direct instruction in small groups begins with a strong teacher involvement that is gradually geared toward peer interaction and peer teaching, with the intent for transfer to each student. The need for early intervention when deficits are evident may be a key to developing independent and metacognitive readers who will be evaluated for mastery of meeting the CCSS. The next chapter outlines the plan for the pilot and full-scale studies that will examine the effectiveness of RT. The information presented in this chapter will be utilized in the preparation and implementation of the study.

Chapter 3 - Methodology

General Overview

This chapter will focus on the quasi-experimental design for the study on the effects of RT as a Tier II intervention for students in the intermediate grades, specifically third, fourth, and fifth grade. IRB approval was obtained for both the pilot and the full-scale study, following permission from the school district to conduct the studies. Since the study involved working with minors, consent forms were sent to students' parents who met the criteria for the experimental or the intervention group. Both the pilot and the full-scale studies were conducted in a suburban school district in the Midwest. The district was close to a major metropolitan area. During the 2016-2017 school year, the district carried a Title I designation based on number of students who qualified for the free and reduced lunch program.

Research Questions

As stated in the purpose of the study, students who struggle to comprehend in the intermediate grades are reading to learn rather than learning to read. Identifying effective and efficient intervention instruction carries not only a professional but also an ethical responsibility to support struggling comprehenders who are to develop into successful learners. This quantitative study is designed to answer:

1. What are the effects of Reciprocal Teaching (RT), as targeted instruction, on standardized literacy assessments for students in the intermediate grades?
2. Are there differences in the effect, if any, RT has on students' standardized test scores at each grade level?
3. Does RT increase students' overall scores on AIMSweb Maze over a period of targeted intervention?

Hypothesis: RT has a significant, positive effect on the performance of intermediate grade level students, post intervention, on standardized comprehension tests.

Null Hypothesis: RT does not have a significant, positive effect on the performance of intermediate grade levels, post intervention, on standardized comprehension tests.

Method

Participants and setting. There were five elementary schools with four sections of each grade level with an average of 20 students per classroom. The district enrollment was 5,835 students K-12 with 2,014 in K-5 elementary schools. The district population, noted in Table 3.1, based on state reporting data, consisted of 52.5% males and 47.5% females; 83.1% Caucasian, 7.5 % Hispanic, 2.2% African American, and 7.2% other; the ELL population was at 1.9% and migrant at 0.2%. A little over 59% were considered low socio-economic status (SES), based on enrollment in the free and reduced lunch program. Students with disabilities were at 21.3%.

Table 3.1 District Demographic Information, 2016-2017

Gender %		Race %				Other %			Students AIMSweb/ Disabilities
<u>Male</u>	<u>Female</u>	<u>African American</u>	<u>Hispanic</u>	<u>Caucasian</u>	<u>Other</u>	<u>ELL</u>	<u>Migrant</u>	<u>Low SES %</u>	
52.5	47.5	2.2	7.5	83.1	7.2	1.9	0.2	59	21.3

Note. Total district enrollment=5,835; elementary enrollment=2,014.

Participant selection procedure. The selection of participants was based on a predetermined by the district cutoff score⁵ on AIMSWeb-Maze comprehension passages for the grade level benchmark. Students scoring below the fall benchmark on AIMSWeb Maze were

⁵ Benchmarks varied between grade levels for AIMSweb Maze (see pilot and full-scale study for specific scores).

invited to participate in the pilot. Students scoring below the 25th percentile on the winter benchmark on AIMSWeb Maze were invited to participate as they were “at-risk” and were the ones considered the most in need of the intervention in the full-scale. Consent forms and an explanation of the study were sent home with the students; in some cases, a follow up phone call or another form of contact was made with the parents/guardians as a reminder to return the forms within a week. After consent forms were signed and returned, the treatment number was determined.

Materials and measurements. AIMSWeb is a standardized assessment by Pearson Education, Inc. (2012) used for screening and progress monitoring in reading and math. The Maze comprehension test consists of cloze-passages where after the first sentence every seventh word is omitted and the student must choose the best word from three given words to go in the blank. The students had three minutes to complete the cloze-passage; only correct responses were counted toward the score (Shinn & Shinn, 2002; Appendix A). Pearson Education, Inc. (2012), reports (a) third grade Maze probes ($\alpha=0.70$, $SEM\ 3.8$); fourth grade Maze probes ($\alpha=0.74$, $SEM\ 3.8$); fifth grade Maze probes ($\alpha=0.78$, $SEM\ 3.9$); (b) a median correlation with end of the year state assessment scores for third grade and Maze scores, $r=0.59$, fourth grade and Maze scores, $r=0.59$, and fifth grade and Maze scores, $r=0.58$; (c) internal-consistency reliability of the Lexile test at each grade ($\alpha=0.90$ to 0.92); (d) significant correlation of AIMSWeb Maze and Lexile levels for third grade, $r=0.58$, for fourth grade, $r=0.58$, and for fifth grade, $r=0.54$. (d) significant correlation of AIMSweb Maze and Lexile levels For 3rd grade Lexile correlation to Maze ($M=7.4$, $SD=9.2$), 4th grade ($M=20.3$; $SD=8.8$), and 5th grade ($M=22.5$, $SD=8.2$) (pp. 14-16). Critical value was not reported; $p<.05$ is assumed.

Experimental variable and general procedure. The experimental variable measured was the effects on students' comprehension scores on standardized literacy assessments, following instruction with Reciprocal Teaching (RT) in grades three through five⁶. The treatment groups at each school received RT as the intervention during small groups, while the control group continued to receive regular instruction in the general education classroom during whole group and other daily small groups. The intervention groups met during Tier II time for 30 minutes, four days per week, in addition to the regular literacy curriculum of 90 minutes. The classroom teachers in each building delivered the small group instruction for the treatment. Each Title I teacher/reading interventionist, observer, and participating classroom teacher attended a half day training session on RT and protocol for delivering the interventions. Teacher trainings that are short in duration and scripted lesson plans replicate what was done by Rinehart et al. (1986).

Descriptions of intervention. The non-fiction texts utilized for the intervention were at the groups' average Lexile level. The non-fiction text was selected from ReadWorks.org (2017), a non-profit online resource. Each student was provided with a folder containing copies of the Article-A-Day text (ReadWorks.org, 2017; Appendix H), which were aligned with CCSS and grade level content, and a graphic organizer with each step of RT: summarizing, questioning, clarifying, and predicting (Appendix M). In addition to the student folders, each classroom teacher and the Title I teacher/interventionist who served as the observer, also received a folder with the planning guide and reproducible graphic organizers, suggested prompts and scaffolding dialogue, and posters/visuals (Appendix D-M). While reading the text using shared reading, the

⁶ Elementary buildings in the district were Pre-Kindergarten/Kindergarten through fifth grade; sixth grade was part of the middle school.

teacher initiated the dialogue after the first paragraph and explicitly modeled while thinking aloud and discussing each part of the graphic organizer. Through a gradual release of responsibility, the students assumed more active role with each subsequent passage as a group and later in peer-teaching. The teacher provided scaffolding, more or less support, for the responses while monitoring and giving feedback, as each student assumed more responsibility for a component of RT (Appendix G). Anchor charts, posters, graphic organizers, and role prompt/cue cards were used to provide support for students during the small group and peer interactions (Appendix I-L).

Research design. The study utilized a Regression Discontinuity (RD) design, which has been used in the past with social policy making and educational decision making, with the standardized assessment serving as the outcome measure (Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007; Jacob & Lefgren, 2004). It was first used by Thistlethwaite & Campbell (1960) on determining criteria for educational merit awards. RD can be used to estimate the effect of an intervention received by those who are determined to “need” it the most (Lesik, 2008; Khandker, Koolwal, & Samad, 2009; Rubin, 2011; Cappelleri and Trochim, 2015). Lesik (2008) cautions to consider the causal effect on the dependent variable as a *probabilistic causality*; in other words that the intervention “*raises the probabilities* of the effect” (positive or negative) it has on the posttest scores for the experimental group (emphasis mine, p. 278). The basic format of the design includes two groups scoring above or below a predetermined cut-off score (Khandker et al., 2009) and a posttest. Two types of RD are possible: sharp and “fuzzy”. In a sharp RD design, the assignment variable is strictly adhered to, as opposed to the “fuzzy” design, and is based on a predetermined rule or criteria (p. 462). (Jacob & Lefgren, 2004; Curs, 2009; Desjardins, McCall, Ott, & Kim, 2010). In order to get a medium effect at 80% power, RD

requires a sample 2.5 times the amount of a random assignment (Laski, 2008). However, adding co-variates alone can increase the power.

Khadker et al., 2009 suggest setting a “narrow band” on each side of the cut-off, from here on out referred to as *caliper* (p. 104). Jacob & Lefgren (2004), discuss that the reason for setting a *caliper*, comparing scores below and above the cut off, is used to show that “unobservable characteristics do not vary discontinuously around the cut off [due to] exogenous variations in the treatment. Because treatment is perfectly correlated with observable characteristics, it is orthogonal, or statistically independent, to unobservable characteristics” (p. 230). Presuming this equality of groups, the effect can be attributed to treatment (intervention in this case) not to differences in groups (Lesik, 2008). In sharp RD, the subjects’ scoring just above and just below the cut-off are likely to be similar and are likely to perform the same prior to intervention based on observable characteristics (Schneider et al., 2007; Khadker et al., 2009; Cohodes & Goodman, 2012). The students in the experimental group are “based on the assignment variable that is exogenous” (Curs, 2008, p. 8; Lesik, 2008). Rubin, 2011 recommends observing for difference in slopes, continuity or discontinuity within the caliper in a sharp design along the regression line. The higher the number of cases around the discontinuity, the higher the precision of the coefficient on the independent variable as the power is increased; the lower the number of cases around the cut off, the less accurate the linear trends (Rubin, 2011).

RD has a strong internal validity due to the selection of participants in the treatment group based on a known cut-off score. Scores are an observable characteristic used in a linear model/function to assign students to a group of 0 or 1 to distinguishing between students above and below the cut-off (Curs, 2009; Rubin, 2011). The regression model is as follow:

$$y = \beta_0 + X\alpha, X \begin{cases} 0 \\ 1 \end{cases}$$

AIMSweb posttest is the outcome variable, β_0 is the treatment effect on the outcome variable, X is the cutoff point for either group⁷--0 or 1, α is the effect on the treatment.

Students are “induced” to participate in the treatment because of their scores; they cannot necessary control the cut-off, but they have otherwise similar characteristics. (Jacob & Lefgren, 2004; Curs, 2009). Desjardins et al. (2010) suggest that RD is close to a randomized experiment because the obscured and unobserved characteristics around the cut off are similar.

One of the fundamental assumptions of RD is that unobservable characteristics vary continuously along the cut-off. It would be hard to manipulate the status of the score even marginally. If there is an impact, there should be discontinuity along the cutoff (Lesik, 2008; Rubin, 2011, WWC, 2010).

The estimated size of the effect of treatment is the discontinuity or difference between the lines of regression at the cut off on the treatment and control groups. Beta (β) represents the treatment effect on the assignment variable (independent variable) or the interaction between the group (control or experiment) and the precut (also referred to as the cut-off score and the dependent variable) (Jacob & Lefgren, 2004, Curs, 2008).

Analysis. In a parametric analysis, RD assumes that (a) cutoff scores are predetermined and strictly adhered to; (b) the distribution is linear; (c) there is variance in the comparison group; (d) comparison and treatment come from the same pre-test; (e) treatment is delivered to all participants equally or the same within the treatment group (WWC, 2010; Cappelleri & Trochim, 2015).

⁷ In this study, the experiment group receiving the intervention is designated as 0 and the control group, not receiving the intervention, is designated as 1.

SPSS was used to run the analysis for all data. Data was analyzed by setting the intercept equal to the cutoff score (see scores for each study) on Maze cloze-passages. After scoring the passages, a group variable was created to distinguish the scores below and above the cut-off (0=below cut off; 1=above cut off). A precut variable was created to represent the distance from the cut off value (pretest-cut score). The values were visually examined on a scatter plot to determine preliminary linear distribution and variance in the group pretest scores.

$$\beta_0 = TREATMENT(TREAT), \beta_1 = CONTROL(CONT)$$

$$\beta_2 PRECUT = (PRETEST - CUTOFF)$$

To make sure that the intervention was the exogenous predictor of scores, a dichotomous variable—interaction, was created between the groups and the precut variable, and its significance examined in relation to the model.

$$\beta_3 INTERACTION(INTRACT) = (GROUP * PRECUT)$$

A univariate ANOVA was used to determine if there was a significant difference between the treatment and control groups. The ANOVA was used in two steps once to determine significant and non-significant variables in the model and another time after removing the previously identified non-significant variables. A variable was considered significant at $p < .05$. After taking the non-significant variables out of the model, parameter estimates were set to show the effects of each predictor and the ANOVA was performed again.

$$MODEL = \beta_0 TREAT + \beta_1 CONT + \beta_2 PRECUT + \beta_3 INTRACT$$

Linear regression analysis followed. An alpha of .05 was used to determine statistical significance; in the event of statistical significance, an effect size was calculated using the standard mean difference (beta, $p < .05$). Scatter plots with regression lines based on the model show the estimated causal effect for each group per grade level.

$$\text{AIMSweb} = \beta * \text{SCORE}(x) + \text{Beta}$$

Validity. The need for early and effective intervention for struggling readers not only in Tier II RtI groups but also in the regular education classroom is addressed through this study. AIMSWeb Maze test administration and scoring followed standard administrative directions and protocol. The selection of the participants in the treatment group was based on a predetermined cut-off score and received no other interventions during Tier II. The large size of the sample was anticipated in order to yield a significant statistical power for the full-scale study; there were 269 students in third, fourth, and fifth grades across three elementary school buildings, 71 qualifying for the intervention group but only 59 students participating in the intervention group who returned signed consent forms. At least three variables were included in the study: the posttests scores (DV), the binary treatment, and the predetermined cut-off score (IV). Attrition was addressed by keeping the groups of students who met the criteria and returned a signed consent form for the duration of the intervention. History was addressed by the three to four-week duration of the intervention. Social threats were addressed by the fact that all students engage in small group work throughout the day. Outside effects, such as reminders for consent to participate in the intervention group, was addressed by teachers contacting the parents either by a typed note, phone calls, or by personal communication (face-to-face or electronic). Treatment related mortality and the end-of-the-school-year anticipation by students was addressed with positive reinforcement tools such as stickers, colorful writing tools, praise, and acknowledgement of effort by the teacher and/or researcher.

Bias and rigor for internal validity was met by the independent variable, which was measured prior to start of the intervention and selection. The cut off score is predetermined months in advance by Pearson Education, Inc. and the school district. All pretest data was

collected prior to the study and students could not have manipulated their scores to qualify for one group or the other. The pilot study, a “fuzzy” RD design, utilized two scores for experimental group participation—Maze and Reading-Curriculum Based Measures (R-CBM) or fluency rate and accuracy. In contrast, the full-scale study had a sharp RD design because it adhered to strict rules for selection of experimental group participants—scores and probes/passages on AIMSweb Maze only, thus controlling for selection bias (Lesik, 2008).

The reading interventionist/Title I teacher utilized a checklist with each item described in the intervention during each session. The researcher conducted observations in each treatment group once per week at each school. Each observation was rated on fidelity on the use of the intervention using the checklist, with the researcher and the reading interventionist/Title I teacher agreeing on 90% of the items described in the intervention procedure. Some lessons were recorded if the interventionist/Title I teachers was not available for observation and scored later.

Pilot Study

Method

Participants and setting. There were five elementary schools with four sections of each grade level with an average of 20 students per classroom. After an initial meeting with administration, one school was selected to participate. Initially, the fourth and fifth grade teams were invited to participate; the fourth-grade team chose to participate in the pilot study. The purpose of the pilot was to examine the overall design of the study. Generalizations or conclusive statements regarding the results of the study were not considered.

Participant selection procedure. The treatment number was determined based on the fall benchmark for AIMSweb Maze ($n < 13$). In order to replicate Palicsar and Brown’s (1983) study, these students also had to be good decoders, or at or above the grade level benchmark on

Reading-Curriculum Based Measures (R-CBM); these are summarized in Table 3.2. After consent forms were signed and returned, the two classrooms, classroom A and classroom B, each had a group for a total of nine students considered “at-risk” that were assigned to the treatment groups; the students above the cut-off were assigned to the control group. Since the purpose of the pilot study was not to form conclusions, the number of students was not an issue for statistical power at the time.

Table 3.2 AIMSweb R-CBM and Maze Testing, Grade 4 Raw Scores

Student	R-CBM ^a		Maze		
	<u>Dec 2016</u>	<u>Jan 2017</u>	<u>Dec 2016</u>	<u>Jan 2017</u>	<u>Feb 2017^b</u>
1	109	123	20	27	38
2	98	127	18	23	38
3	134	146	16	24	26
4	110	113	14	17	29
5	129	155	17	30	27
6	118	127	19	27	18
7	108	147	23	29	23
8	131	139	10	20	21
9	105	109	18	29	20

Note. ^a R-CBM represents the scores on a Reading-Curriculum Based Measures indicating the fluency of decoding, specifically accuracy and rate or words read correct in one minute (WCPM). ^b Scores for testing one week after intervention was discontinued in one classroom (students 5-9) but not in the other (students 1-4).

The demographic information for each group in classroom A and classroom B is presented in Table 3.3. Classroom A consisted of four students, one female and three males; all students were Caucasian. Classroom B consisted of five students, all males; one student was African American and four were Caucasian. As reported, none of the students in classroom A and classroom B participated in special education services.

Table 3.3 Fourth-grade students participating in the RT intervention

<u>Gender</u>		<u>Race/Ethnicity</u>				<u>Low SES^a</u>		<u>SPED^b</u>	<u>Mean Age^c</u>
Male	Female	African American	Asian	Caucasian	Hispanic	Free	Reduced		
8	1	0	0	8	1	2	1	0	9.9

Note. ^a Low Socio-Economic Status (SES) is based on free and reduced lunches. ^b SPED may include learning, intellectual or physical disability; behavior, autism spectrum, or communication disorder. ^c Age is presented in years and months.

Materials and measurements. The Maze comprehension test consists of cloze-passages where after the first sentence every seventh word is omitted and the fourth-grade students must choose the best word from three given words to go in the blank. The students had three minutes to complete the cloze-passage; only correct responses are counted toward the score. Pearson Education, Inc. (2014), reports (a) significant reliability among fourth grade Maze probes ($\alpha=0.74$, $SEM\ 3.8$); (b) a median correlation with end of the year state assessment scores for fourth grade and Maze scores, $r=0.59$ ($Mdn=0.57$); (c) internal-consistency reliability of the Lexile test at each grade ($\alpha=0.90$ to 0.92); (d) significant correlation of AIMSweb Maze and Lexile levels for fourth grade, $r=.58$. Critical value was not reported; $p<.05$ is assumed.

Experimental variable and general procedure. For the pilot study, the fall 2016 benchmark was used to determine eligibility for the intervention group. The students took the test in late August of 2016, which provided the pre-test score or the set cut-off score. The winter benchmark test was given in the last full school week of December and a progress monitoring probe was given as the post-test score in the pilot study. The intervention lasted for approximately five weeks, two weeks in December and three weeks in January. The classroom teachers in each classroom delivered the small group instruction for the treatment. The procedure described in the general overview section was applied to the pilot study.

Research design. The pilot study utilized a Regression Discontinuity (RD) design. The cut-off score for the treatment was the AIMSweb Maze fourth grade benchmark for fall ($n>13$); the posttest target was also set according to AIMSweb Maze for the winter benchmark ($n>19$). The pilot study utilized the same parametric RD procedure as described in the overview section for analysis and validity measures.

Full-Scale Study

Method

Participants and setting. There were five elementary schools with four sections of each grade level with an average of 20 students per classroom. All schools were invited to participate; only the teachers who were interested participated in the full-scale study.

Participant selection procedure. After consent forms were signed and returned, the treatment number was determined (N=59), for an alpha .05. It consisted of 54% boys and 45% girls; 12% African American, 0% Asian, 83% Caucasian, 2% Hispanic, and 3% other. For each school, students considered “at-risk”, below the 25th percentile on the winter AIMSweb testing were assigned to treatment groups of four to six; the students above the cutoff were assigned to the control group.

Table 3.4 Full-Scale Study Student Demographic Information

<u>Grade</u>	<u>Gender</u>		<u>Race/Ethnicity</u>					<u>Low SES^b</u>
	Male	Female	African American	Asian	Caucasian	Hispanic	Other	Title I/Title I Identified
Third	15	18	4	0	32	0	1	Yes
Fourth	7	7	2	0	12	0	0	Yes
Fifth	10	2	1	0	5	1	1	Yes
Total	32	27	7	0	49	1	2	N=59

Note. ^a Students who met the criteria for participation in the intervention group who returned a signed consent form. ^b Based on Free and Reduced lunches—aggregate data only. SPED status and Mean Age were not provided.

Materials and measurements. Pearson Education, Inc. (2012), reports (a) significant reliability among third grade Maze probes ($\alpha=0.70$, *SEM* 3.8); fourth grade Maze probes ($\alpha=0.74$, *SEM* 3.8); fifth grade Maze probes ($\alpha=0.78$, *SEM* 3.9); (b) a median correlation with

end of the year state assessment scores for third grade and Maze scores, $r=0.59$, fourth grade and Maze scores, $r=0.59$, and fifth grade and Maze scores, $r=0.58$; (c) internal-consistency reliability of the Lexile test at each grade ($\alpha=0.90$ to 0.92); (d) significant correlation of AIMSweb Maze and Lexile levels for third grade, $r=0.58$, for fourth grade, $r=0.58$, and for fifth grade, $r=0.54$. Critical value was not reported; $p<.05$ is assumed. For third grade Lexile correlation to Maze was ($M=7.4$, $SD=9.2$), for fourth grade ($M=20.3$; $SD=8.8$), and fifth grade ($M=22.5$, $SD=8.2$) (pp. 14-16).

Lesson plans were provided to teachers for the week, including the non-fiction text on the group's average Lexile level (Appendix H). Each lesson plan followed the same protocol of before, during, and after reading; it included direct instruction model consisting of modeling, guided practice, gradual release of responsibility, and scaffolding. When selecting non-fiction text for the groups, students' most common interests were considered. Students' most common interests were determined by using Johns and Lenski's (2014) interest inventory, *News About Me* (Appendix F). An Excel sheet was used to find the frequency of interests and the top three for each classroom or grade level. Some of the more common interests included animals, collecting rocks, friendships, careers, and favorite school subject.

Experimental variable and general procedure. The winter test was given in the last full school week of December; these scores were used to determine the participant eligibility, students with scores below the 25th percentile on AIMSweb Maze, for each grade level. Approximately two weeks were allowed for consent forms to be returned. The intervention lasted for four weeks, except for three 3rd grade classrooms, where the intervention lasted for three weeks due to other testing scheduling. The test for the spring benchmark was given at the end of the second week in May. The classroom teachers in each building delivered the small group

instruction for the treatment. Besides using the cut-off score, pre- and post-tests were implemented using Maze passages to determine if the intervention was the cause of change in scores; this was based on the suggestion by Jacob & Paris' (1987) study, where the authors did not implement a pre-post measure immediately before and after RT was applied. The intervention for the full-scale study followed the same procedures as described in the overview section.

Research design. RD design, as described in the overview section, was also applied for the full-scale study. The scores were obtained from the AIMSweb Maze winter with the following benchmark for third (≥ 14), fourth (≥ 19), and fifth grade (≥ 21); the posttest target was also set according to AIMSweb Maze for the spring benchmark for third (≥ 15), fourth (≥ 19), and fifth grade (≥ 25).

The participant selection for the study, based on RD, were students who scored below the 25th percentile on AIMSweb's Maze winter comprehension cloze passages (Appendix A); the cut-off (raw) scores were 10 for third, 14 for fourth, and 16 for fifth. To replicate Palincsar & Brown's (1983, 1984) study, these students also had to be good decoders. However, in order to avoid a mis-assignment of participants and to sustain a sharp RD design (Capelleri & Trochim, 2015), only students below the 25th percentile cut-off on comprehension on AIMSweb's Maze winter were selected for the intervention. The text was controlled for decoding based on the groups' average Lexile level.

A caliper was set below and above the cut-off score, where the focus was on students who were similar in scores along the cut-off on AIMSweb Maze for comprehension only; randomization is assumed as students are presumed to be statistically similar based on observable characteristics (Schneider et al., 2007; Khadker et al., 2009; Cohodes & Goodman,

2012). This would allow for a “sharp” RD design with clear boundaries for comparison of participants with similar observable characteristics and the effect of RT on their comprehension scores of pre/post testing. Four units below and above the cut score served as the caliper or the boundaries around the cut off score. The decision to set this caliper around the score was based on SEM on the reported tests stability per grade level over a four-month period, fall to winter for Maze (Pearson Education, Inc., 2012, p. 15). The variations for the raw scores for each grade level from the mean are listed as following in the manual: third reliability of 0.70, SEM=3.8; fourth 0.74, SEM=3.8; and fifth of 0.79, SEM=3.9; the caliper was set at 4 points below and above the cut-off for each grade sample. The full-scale study utilized the same parametric RD procedure as described in the overview section for analysis and validity measures.

Summary

This chapter discussed the procedure, setting, participants, and intervention protocol. Common parts for both studies were included in the overview, after which specifics were detailed in one section for the pilot study and another section for the full-scale study. In this study, the RD design is applied to differentiate between the control and experimental group by a predetermined cut-off score. In the pilot study, a small sample was used and all scores were considered in the regression model. For the full-scale study, a larger sample was secured and students’ scores below and above the cut-off were included as well. Another step was added to the full-scale study—looking at scores within a caliper that were closely surrounding the cut-off. According to RD, students performing just above and below the cut-off point would typically perform in a similar way given the same test over time. By concentrating on the students’ scores within the caliper allows for comparison of groups of the students with similar observable characteristics and more precise explanation for the effect of the intervention can be developed.

Chapter 4 - Results

The Sights and Sounds During Intervention

Overview of daily lesson plans. The procedure for the pilot and the full-scale study followed a trend of lesson progression. The teachers first introduced each strategy one at a time with explicit modeling and thinking aloud of how they apply it with the text (Appendix B). The teacher also showed kinesthetic model of each strategy that tied in with the poster or anchor charts symbol. For example, predicting was depicted by a crystal ball and the action for this was the motion of hands in the shape of a ball; questioning was a microphone, pretending to be a reporter who “interviews” the group with a question. Clarifying was a time out sign, as in a sports game, with the hands making a T-sign; for summarizing, the students pretended to use a reel camera, reminding them to pick out the most important events as if retelling a movie.

During the next sequence of lessons, the teachers modeled the strategies together, as they read aloud and the students only followed along in the text (Appendix C). Once modeled, students had the opportunity to practice what was modeled and the teacher monitored to provide feedback and appropriate scaffolding, and to gauge the amount gradual release. At the beginning of each intervention meeting, the teachers reviewed each strategy and continued to provide feedback and appropriate scaffolding and gradual release. In the pilot, the Title I teacher and the researcher alternated observations in each classroom since the fourth-grade reading intervention occurred at the same time (Appendix E); during the full-scale, a schedule of observations with the Title I teacher or the interventionist, the teachers, and the researcher was shared in Google drive to coordinate observations in each classroom, once per week. The teachers provided input and feedback for weekly lesson planning and for copies of more text from ReadWorks.org via e-mail or verbally when checking in with the researcher at observations.

The teachers post-tested all students on the Maze passages in mid January--those participating in the intervention groups and the rest of the fourth graders in both classes in the pilot study. The intervention continued in one classroom for one more week but it was discontinued in the other; all students in each intervention group were tested again as a progress monitor measure a week later. In the full-scale study, all teachers pre- and post-tested the students in the intervention groups, immediately before and after the intervention. All students, including the ones in the intervention group, in grades three, four, and five were given a spring benchmark Maze passages. Results for both studies were shared with administrators, the teachers, and the Title I teacher/interventionists.

Lesson sequencing and organization. Lesson plans, graphic organizers, and prompts were either used or adapted for classroom use only from *Reciprocal Teaching at Work* (Oszkus, 2010). The weekly lesson plans included non-fiction, expository/informational reading passages at the group's average Lexile levels from the ReadWorks.org (2017) website as hard copies and in Google files for easy sharing with the teachers; for the full-scale study, the files were organized by school, grade level, and week. Each teacher binder included copies of graphic organizers for students to write in, role cards for Predictor, Questioner, Clarifier, and a Summarizer, games, response prompts, and anchor charts for each strategy.

The lesson plans were organized in a Before, During, and After (BDA) sequence, which was delivered using direct instruction. As seen in Appendix C, each lesson plan included the objective, which was shared with the group of students and an anticipatory set was used to pique the students' interest and activate their schemata about the topic from the article. The anticipatory set was either a question about the topic or sharing a story relative to the topic that students could connect to. The teachers first modeled each strategy with a short text or a

paragraph, while still allowing some student dialogue to occur. This served as an informal assessment to decide if scaffolding was needed, who needed it, and how much gradual release the teacher could allow for students to work in pairs and peer teach.

The teachers modeled making predictions about the topic by focusing on text features, previewing the content-specific vocabulary, or scanning thorough the text—all previously used methods in the classroom from other curriculum resources. Questioning and clarifying occurred when students were not as familiar with the topic or with content-specific vocabulary. The groups used previously taught methods for question generation and for clarifying ideas or words, such as the 5Ws & How (who, what, when, where, why, and how), looking for context clues, or rereading. Summarizing, typically used at the end of the text or a paragraph, required much more modeling and scaffolding. Teachers and students recalled what they had been using throughout the year for developing a summary. Teachers had to monitor closely to ascertain that students' responses were text based, not schema based and to provide corrective feedback when and if this occurred. If this occurred, teachers reminded students to go back to the text and find evidence to support their prediction, question, clarification, or summary and/or model this for the students.

Gradually, through informal assessments, the teachers assigned “roles” for each student in the group. Each student had a “role” card, folded in half, in front of them; the role, like Questioner was printed on the front, and the prompts reminding them of their “job”, on the back. The teachers took on a more passive role but were still available to provide support and feedback to students, to reaffirm their effective use of the strategies, and to praise their efforts. By the last week in both stages of the study, the teachers in the classrooms had pairs of students, peer teaching, who utilized the four strategies with non-fiction text and with little to no support.

Classroom setting. The students were in groups between four and six and they met at a horseshoe table with the teacher during the intervention time. The intervention took place during the Tier II time at the grade level either on Mondays, Tuesdays, Wednesdays, and Thursdays or on Tuesdays, Wednesdays, Thursdays, and Fridays for thirty minutes. Students quickly adapted to the routine of getting their folders and other necessary materials to the horseshoe table since small groups were utilized throughout the day in the elementary buildings. While the intervention group worked with the teachers, the rest of the class was either in other small groups lead by other adults in the classroom or in the hallway, or working independently at their desk or in the hallway with adult supervision.

When students worked in pairs, they were away from the horseshoe table; they were either at another table or desk, on the rug where classroom meetings take place, or another location of their choice in the classroom. They had their folders with the text, along with their choice of highlighters, sticky notes, and a pen or a pencil. When they were finished, the students conferenced with the teacher on each strategy after recording information on their graphic organizers. The teacher could scaffold students' understanding and interpretation of the text for accuracy, providing feedback, prompting and guiding correct responses, and asking students to go back and look in the text for evidence. At the point of peer teaching, the students had more independence with strategic reading but still had the support of the teacher when finished with the reading of the text.

Pilot Study

The increase in scores and the overall feelings about the study were positive, even though the selection based on the design proved to have some mis-assignment of participants based on two qualifying scores (see Table 4.2). There were 32 students in the control group of fourth graders. The sample for the pilot consisted of eight males and one female for a total of nine fourth-graders (N=9), of which eight were Caucasian and one Hispanic students (see Table 4.1). The nine students participated in two intervention groups in classrooms A and B. The group in classroom A was led by the classroom teacher and consisted of three males and one female; the group in classroom B was led by the classroom teacher and consisted of five males.

Table 4.1 Pilot Study Student Demographic Information

Fourth-grade students participating in the RT intervention

<u>Gender</u>		<u>Race/Ethnicity</u>				<u>LoW SES^a</u>		<u>SPED^b</u>	<u>Mean</u>
Male	Female	African American	Asian	Caucasian	Hispanic	Free	Reduced		<u>Age^c</u>
8	1	0	0	8	1	2	1	0	9.9

Note. ^a Low Socio-Economic Status (SES) is based on free and reduced lunches. ^b SPEDV may include learning, intellectual or physical disability; behavior, autism spectrum, or communication disorder. ^c Age is presented in years and months.

The fall 2016 AIMSweb Maze benchmark was used for the pre-test and the selection for participation in the intervention (M=25.22, SD=4.46). One week before the end of the study, the students were post tested with the winter 2017 benchmark in both classrooms. One classroom continued the intervention and the other did not. There was an overall increase in raw scores (M=26.67, SD=6.14). Students one through four, and student eight had an increase in scores from pre-test to post-test. Students five through seven, and nine show a decline in scores; they were also in the group where the intervention lasted a week less than the other classroom.

Table 4.2 AIMSweb R-CBM and Maze Testing, Grade 4 Raw Scores

Student	R-CBM ^a		Maze		
	<u>Dec 2016</u>	<u>Jan 2017</u>	<u>Dec</u> <u>2016</u>	<u>Jan 2017</u>	<u>Feb 2017^b</u>
1	109	123	20	27	38
2	98	127	18	23	38
3	134	146	16	24	26
4	110	113	14	17	29
5	129	155	17	30	27
6	118	127	19	27	18
7	108	147	23	29	23
8	131	139	10	20	21
9	105	109	18	29	20

Note. ^a R-CBM represents the scores on a Reading-Curriculum Based Measures indicating the fluency of decoding, specifically accuracy and rate or words read correct in one minute (WCPM). ^b Scores for testing one week after intervention was discontinued in one classroom (students 5-9) but not in the other (students 1-4).

A one-way between subjects ANOVA was conducted to compare the effect of the intervention on the post-test scores. There was a significant effect of the intervention on the posttest at the $p < .05$ level for the experiment group [$F(1, 16) = 35.08, p = 0.00$].

The effects of the intervention (RT) on the post-test scores were statistically significant ($\beta = 0.828, p < 0.05$) for the experiment group but not statistically significant for the control ($\beta = -0.207, p < .05$). In other words, for every unit of the independent variable—scores below the cut off, the dependent variable—post scores on AIMSweb Maze, increased by 1.089 points for the experiment group. Because students were not tested immediately before and immediately after the intervention (within 24 hours), it is unknown whether variables, other than the intervention, may have influenced the scores. Further investigation with a larger sample was needed to determine a more statistically sound explanation for the results of this pilot study.

Table 4.3 Effects on Post-Test Scores

Group	<i>B</i>	<i>SE B</i>	β	<i>t</i>	<i>p</i>
Intervention	1.098	0.186	0.828	5.918	0.000
Control	-0.322	0.349	-0.207	-0.922	0.368

* $p < 0.05$, two tailed test

The daily lesson plans followed the same protocol for each text. The observers, the Title I teacher and the researcher, agreed on average of 91.5% of the items in the first two weeks, and 86.5% in the last two weeks. With the goal for inter-observer agreement being 90%, one possible explanation would be more release of responsibility to the students, from teacher-guided to peer-guided instruction, and less modeling.

Full-Scale Study

After completing the analysis of the data for all students, RT had a positive effect as a targeted instruction on standardized literacy assessments for students in the intermediate grades in the present study. The effect was not statistically significant for third ($\beta_3=0.017$, $p > .05$) and fourth ($\beta_4=0.442$, $p > .05$) grades, but it was statistically significant for fifth grade ($\beta_5=0.767$, $p < .05$). The analysis was first performed for all students in the experimental group, students below the 25th percentile on their winter testing. The results, when considering all of the students at each grade level, indicated that for every unit of the independent variable—pre-test scores below the cut-off, the dependent variable—post scores on AIMSweb Maze for fifth grade increased by 1.17 points; there was an insignificant effect in points increase for third (0.027) and fourth (0.476) grades. A one-way between subjects ANOVA was conducted to compare the effect of the intervention on the post-test scores for each grade level and is discussed in the grade level sections below.

Table 4.4 Full-Scale Study Student Demographic InformationStudents participating in the RT Intervention ^a

<u>Grade</u>	<u>Gender</u>		<u>Race/Ethnicity</u>					<u>Low SES^b</u>
	Male	Female	African American	Asian	Caucasian	Hispanic	Other	Title I I/Title I Identified
Third	15	18	4	0	32	0	1	Yes
Fourth	7	7	2	0	12	0	0	Yes
Fifth	10	2	1	0	5	1	2	Yes
Total	32	27	7	0	49	1	2	N=59

Note. ^a Students who met the criteria for participation in the intervention group who returned a signed consent form. ^b Based on Free and Reduced lunches—aggregate data only. SPED status and Mean Age were not provided.

A net effect (β_{nA}) of all scores for all grades combined was not statistically significant ($p < 0.391$). This is possibly due to the pre-established, varying benchmark goals for each grade level that create differences in the percentile cut-off scores. A one-way between subjects ANOVA was conducted to compare the effect of the intervention on the post-test scores. There was a significant effect of the intervention on the posttest at the $p < 0.05$ level for the experiment group for one of the three grade levels: third grade [$F(1, 52) = 0.016, p = 0.901$]; fourth grade [$F(1, 14) = 3.403, p = 0.086$]; and fifth grade [$F(1, 28) = 38.533, p = 0.00$]. For all third and fourth graders in the experimental group, RT was not statistically significant but it was statistically significant for the fifth graders in the experimental group. RT also had a positive effect on the performance of intermediate grade level students, post intervention, on standardized

comprehension tests. In this case, there was a fail to reject the null for the third and fourth grades. The hypothesis was accepted for fifth grade, and the null was rejected.

Table 4.5 AIMSweb Maze, Grades 3-5

Benchmark Testing—All Students				
<u>Third Grade</u>	<u>N</u>	<u>M</u>	<u>MSE</u>	<u>SD</u>
Winter	145	12.48	.562	6.767
Spring	133	13.69	.611	7.043
Valid N (listwise)	130			
<u>Fourth Grade</u>				
Winter	38	16.97	1.488	9.172
Spring	38	17.89	1.817	11.200
Valid N (listwise)	38			
<u>Fifth Grade</u>				
Winter	101	22.65	1.000	10.047
Spring	102	26.95	.943	9.525
Valid N (listwise)	101			

After completing the analysis of the data with a set caliper for the scores around the cut-off ($PRECUT \geq -4$ AND ≤ 4), RT had a positive but statistically insignificant effect as a targeted instruction on standardized literacy assessments for students in the intermediate grades, who are assumed to be similar in ability based on observable characteristics. This was indicated in the present study for third ($\beta_3=0.172$, $p > .05$), fourth ($\beta_4=0.198$, $p > .05$), and fifth ($\beta_5=0.327$, $p > .05$) grade. A one-way between subjects ANOVA was conducted to compare the effect of the intervention on the post-test scores. A net (β_{nc}) effect of scores within the caliper for all grades was not statistically significant ($p < 0.826$). As with the net effect outside of the caliper, this is possibly due to the pre-established differences in the cut-off scores and varying benchmark goals for each grade level. For within the caliper scores, there was a statistically insignificant effect of the intervention on the posttest at the $p < 0.05$ level for the experiment group for all grade levels:

third grade [$F(1, 39) = 1.183, p = 0.283$]; fourth grade [$F(1, 6) = 0.246, p = 0.638$]; and fifth grade [$F(1, 13) = 1.551, p = 0.235$]. This relates to research questions one and three, in which case the hypothesis that RT has a significant, positive effect on the performance of intermediate grade level students, post intervention, on standardized comprehension tests cannot be accepted; however, there was a fail to reject the null. For every one-point increase in pretest scores, the post scores on AIMSweb Maze increased by 0.457 of a point for third grade, 0.780 of a point for fourth, and 0.685 of a point for fifth grade, which makes them statistically insignificant. In this case RT also had a positive effect on the performance of intermediate grade level students whose scores were close to the cutoff by a margin of ± 4 , post intervention, on standardized comprehension tests. The hypothesis was not accepted and there was a fail to reject the null. The results are statistically insignificant, as the increase is not substantiated by even one point.

Table 4.6 AIMSweb Maze, Grades 3-5

Benchmark Testing—Around Cut-Off ± 4				
<u>Third Grade</u>	<u>N</u>	<u>M</u>	<u>MSE</u>	<u>SD</u>
Winter	82	9.90	0.279	2.522
Spring	70	10.13	0.474	3.963
Valid N (listwise)	70			
<u>Fourth Grade</u>				
Winter	15	14.27	0.573	2.219
Spring	15	13.33	1.709	6.619
Valid N (listwise)	15			
<u>Fifth Grade</u>				
Winter	31	16.29	0.525	2.923
Spring	31	21.84	0.674	3.751
Valid N (listwise)	31			

Effects per grade level within the caliper

Third grade. Third grades from three elementary schools were invited to consider

participating in the study; there was one section of third grade from School 1, two sections in School 2, and three sections in School 3. The total number of third graders between all the schools were 129⁸; the control group included 116 students and the experiment group participating in the intervention included 35 third-graders after consent forms were returned.

The winter 2017 benchmark was used for the pre-test and the selection for participation in the intervention of students below the 25th percentile on AIMSweb Maze (M=9.90, SD=2.522). The students were post tested with the spring 2017 benchmark, showing a slight increase in average raw scores (M=10.13, SD=3.963).

There is a pattern worth noting that emerged when comparing the pre- and post-score gains immediately before and after the intervention, with the scores on the winter and spring benchmark gains (see Table 4.7). For the pre/post immediately before and after 18 students received negative gain scores, and three students had no change, or a gain score of 0. Student 13 and 20 were not pre/post-tested immediately before and after, and student 24 and 31 were not post-tested immediately before and after. The winter and fall benchmark gain scores, show that eight students had negative gain scores, and three students had no change, or a gain score of 0. Positive gains from the intervention are seen for two students (4 and 9), while the rest of the third- grade students had negative gains when comparing pre/post intervention and winter/spring benchmark gain scores. These results may suggest that the positive gains in student scores may not be due to the intervention. However, when looking at the post test scores and the spring benchmark score, seven students scored higher on the spring benchmark than they did on the post-test immediately following the intervention (11, 12, 17, 22, 26, 27, and 29), one student

⁸ (+3) students at post for 3rd grade, School 1; (-17) students at post for 3rd grade, Schools 2 and 3 or (143-14)

had no change (33), while 22 third-graders scored higher immediately following the intervention than they did on the spring benchmark assessment.

Table 4.7 Third Grade Gain Scores

Immediate Pre/Post-Testing & Benchmark

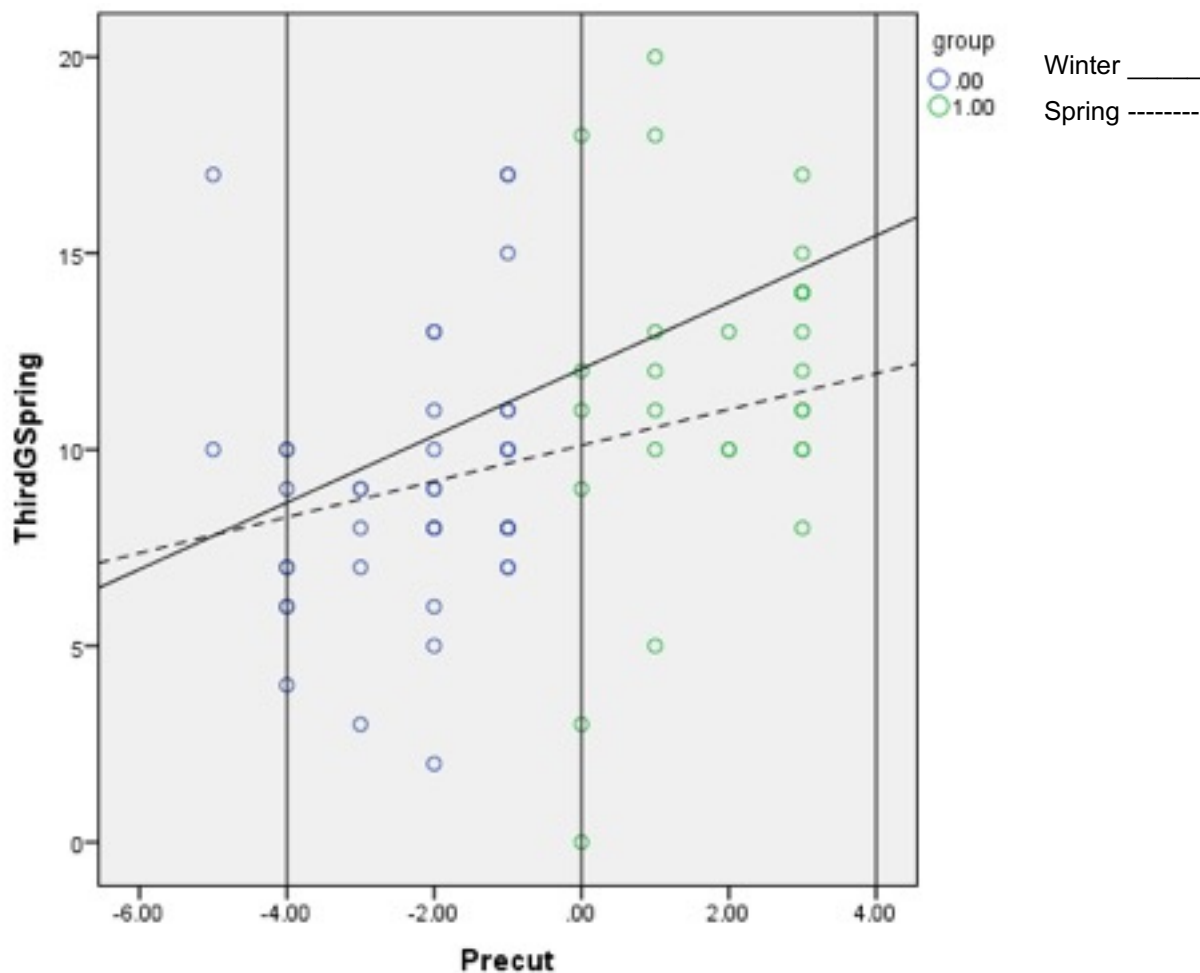
<u>Grade 3</u>	<u>Pre-Test</u>	<u>Post-Test</u>	<u>Gain Score</u>	<u>Winter Benchmark≥ 14</u>	<u>Spring Benchmark≥ 15</u>	<u>Gain Score</u>
1	11	12	1	10	11	1
2	16	16	0	10	15	5
3	13	18	5	10	17	7
4	16	18	2	10	8	-2
5	18	24	6	9	13	4
6	15	20	5	7	9	2
7	5	12	7	6	10	4
8	13	10	-3	10	7	-3
9	12	14	2	10	8	-2
10	11	8	-3	7	6	-1
11	12	14	2	9	17	8
12	18	15	-3	9	16	7
13	--	--	--	9	8	-1
14	16	15	-1	7	10	3
15	14	14	0	5	10	5
16	17	15	-2	4	10	6
17	13	7	-6	4	8	4
18	10	9	-1	3	7	4
19	14	15	1	10	10	0
20	--	--	--	8	8	0
21	19	15	-4	7	14	7
22	16	14	-2	7	17	10
23	19	15	-4	7	10	3
24	5	--	--	6	4	-2
25	17	12	-5	4	10	6
26	9	6	-3	4	7	3
27	9	6	-3	2	7	5
28	10	9	-1	2	5	3
29	18	13	-5	10	15	5
30	23	27	4	10	19	9
31	12	n/a	--	10	9	-1
32	12	9	-3	9	7	-2
33	15	15	0	8	8	0
34	13	11	-2	5	11	6
35	11	10	-1	5	13	8

Note. An average gain score of pre/post 3.53 points; benchmark gain of 5.66.

Thirteen students' post scores immediately following the intervention were at or above the spring benchmark of 15, while only seven students remained at or above the spring benchmark at the time of benchmark testing.

The regression discontinuity analysis showed that an insignificant, positive main effect ($\beta_3=0.172$, $p > .05$), was observed for the total standard score for the intervention ($GROUP=0$). Regression discontinuity analysis between groups indicated that an insignificant main effect existed for the control ($GROUP=1$), ($b=0.257$, $p > .05$).

Figure 4.1. Third Grade Discontinuation at Cut-Off Point Comparing Pre-Test Group Scores and Post-Test Group Scores Following Treatment.



Fourth grade. Fourth grades from three elementary schools were invited to consider participating in the study; there were two sections of fourth grade from School 3. The total number of fourth-graders between the two sections in the school was 38; the control group included 26 students and the experiment group participating in the intervention included 12 fourth graders after consent forms were returned.

The winter 2017 benchmark used for the pre-test and the selection for participation in the intervention of students below the 25th percentile on AIMSweb Maze (M=14.27, SD= 2.219). The students were post-tested with the spring 2017 benchmark, showing a slight decrease in average raw scores (M=13.33, SD= 6.619).

Table 4.8 Fourth Grade Gain Scores

Immediate Pre/Post-Testing & Benchmark

<u>Grade 4 & Student #</u>	<u>Pre-Test</u>	<u>Post-Test</u>	<u>Gain Score</u>	<u>Winter Benchmark\geq19</u>	<u>Spring Benchmark\geq19</u>	<u>Gain Score</u>
1	20	16	-4	14	14	0
2	15	14	-1	10	10	0
3	20	16	-4	7	6	-1
4	5	8	3	7	8	1
5	5	11	6	6	6	0
6	12	15	3	4	8	4
7	--	5	--	2	6	4
8	--	--	--	14	--	--
9	--	--	--	11	--	--
10	4	7	3	0	--	--
11	13	18	5	9	11	2
12	13	12	-1	8	9	1

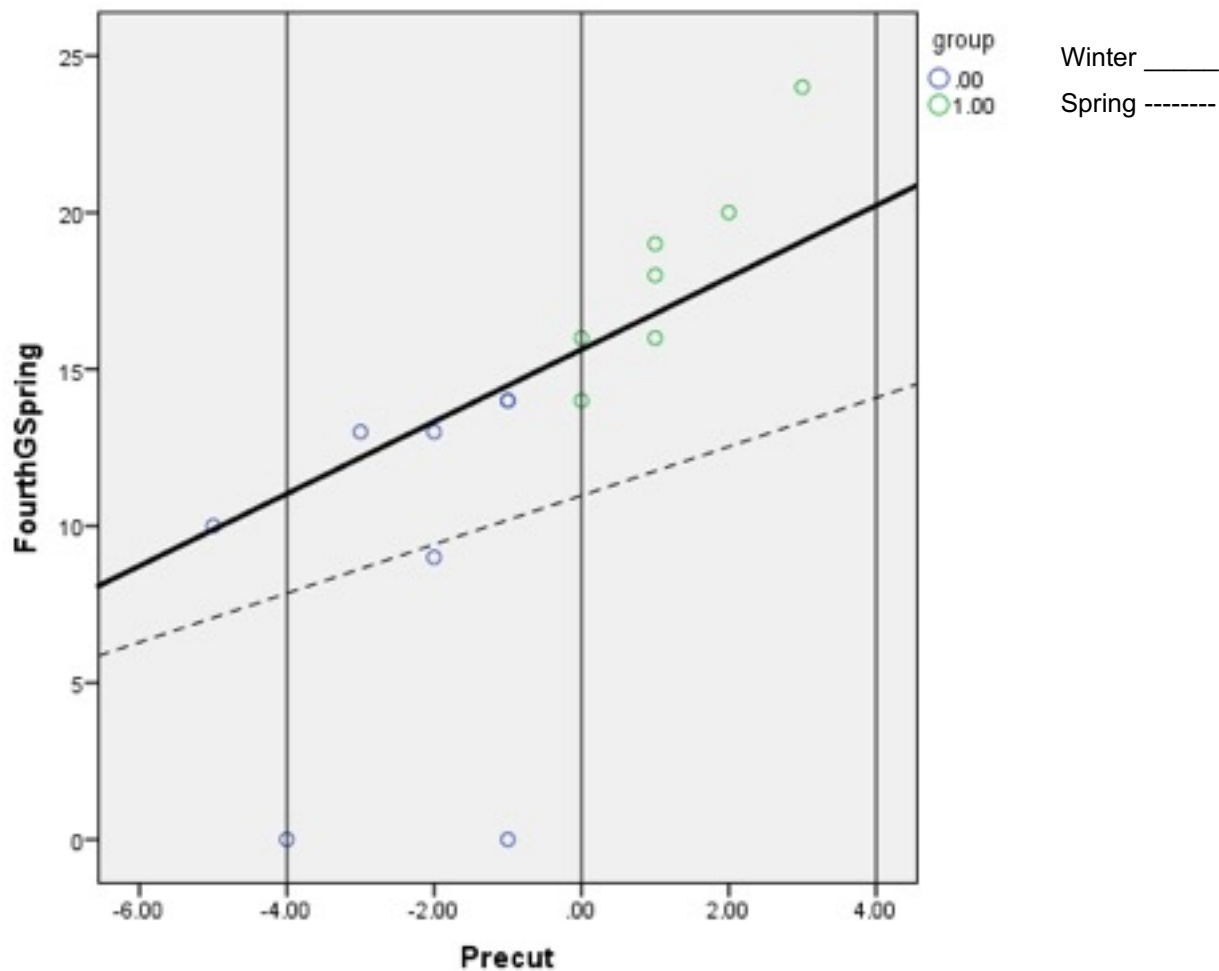
Note. An average gain score of pre/post 2.33 points; benchmark gain of 2.33.

For the pre/post immediately before and after, four students received negative gain scores, and five students had positive gains. Students 7, 8, and 9 were not pre/post-tested immediately before and after; students 8, 9, and 10 were not post-tested for the spring

benchmark. The winter and fall benchmark gain scores, show that one student had negative gain scores, and three students had no change, or a gain score of 0. Positive gains from the intervention are seen for three students (4, 5, and 11), while five of the fourth-grade students had negative gains when comparing pre/post intervention and winter/spring benchmark gain scores; due to the missing scores for four of the fourth-grades, it is unclear as to whether there were negative or positive gains for students seven through ten. When comparing the post test scores and the spring benchmark score, one student scored higher on the spring benchmark than they did on the post-test immediately following the intervention (7), one student had no change (4), while the rest of the seven fourth-graders scored higher immediately following the intervention than they did on the spring benchmark assessment. Three students (8, 9, and 10) had missing scores that could not be compared. All students' scores remained below the spring benchmark of 19, regardless whether they were tested immediately after the intervention or at spring benchmark testing.

The regression discontinuity analysis showed that an insignificant, positive main effect ($\beta_4=0.198$, $p>.05$), was observed for the total standard score for the intervention ($GROUP=0$). Regression discontinuity analysis between groups indicated that a significant main effect existed for the control ($GROUP=1$), ($b=0.942$, $p<.05$).

Figure 4.2. Fourth Grade Discontinuation at Cut-Off Point Comparing Pre-Test Group Scores and Post-Test Group Scores Following Treatment.



Fifth grade. Fifth grades from three elementary schools were invited to consider participating in the study; there was one section of third grade from School 1, one section from School 2, and one section in School 3. The total number of fifth graders between all the schools was 102; the control group included 89 students and the experiment group participating in the intervention included 12 fifth-graders after consent forms were returned.

The winter 2017 benchmark used for the pre-test and the selection for participation in the intervention of students below the 25th percentile on AIMSweb Maze ($M=16.29$, $SD=2.923$). The

students were post tested with the spring 2017 benchmark, showing an increase in average raw scores ($M=21.84$, $SD=3.751$).

Table 4.9 Fifth Grade Gain Scores

Immediate Pre/Post-Testing & Benchmark

<u>Grade 5 &</u> <u>Student #</u>	<u>Pre-</u> <u>Test</u>	<u>Post-</u> <u>Test</u>	<u>Gain</u> <u>Score</u>	<u>Winter</u> <u>Benchmark\geq21</u>	<u>Spring</u> <u>Benchmark\geq25</u>	<u>Gain</u> <u>Score</u>
1	20	24	2	15	21	6
2	21	27	6	14	19	5
3	23	18	-5	12	24	12
4	14	12	-2	10	14	4
5	14	17	3	12	17	5
6	5	14	9	7	9	2
7	18	25	7	11	12	1
8	14	30	16	16	29	13
9	13	16	3	16	19	3
10	18	16	-2	15	18	3
11	18	19	1	15	19	4
12	14	13	-1	6	12	6

Note. An average gain score of pre/post 7.85 points; benchmark gain of 16.67.

For the pre/post immediately before and after, four students received negative gain scores, and eight students had positive gains. The winter and spring benchmark gain scores, show that none of student had negative gain scores. Positive gains from the intervention are seen for four students (2, 6, 7, and 8), while the rest of the fifth-grade students had negative gains when comparing pre/post intervention and winter/spring benchmark gain scores. For the post test scores and the spring benchmark score comparison, four students scored higher on the spring benchmark than they did on the post-test immediately following the intervention (3, 4, 9, and 10) two student had no change (5 and 11), while the rest of the six fifth-graders scored higher immediately following the intervention than they did on the spring benchmark assessment. Three students' post scores immediately following the intervention were at or above the spring

benchmark of 25, while only one student remained at or above the spring benchmark at the time of benchmark testing.

The regression discontinuity analysis showed that an insignificant, positive main effect ($\beta_5=0.327$, $p>.05$) was observed for the total standard score for the intervention ($GROUP=0$). Regression discontinuity analysis between groups indicated that an insignificant main effect existed for the control ($GROUP=1$), ($b=0.347$, $p>.05$).

Figure 4.3. Fifth Grade Discontinuation at Cut-Off Point Comparing Pre-Test Group Scores and Post-Test Group Scores Following Treatment.

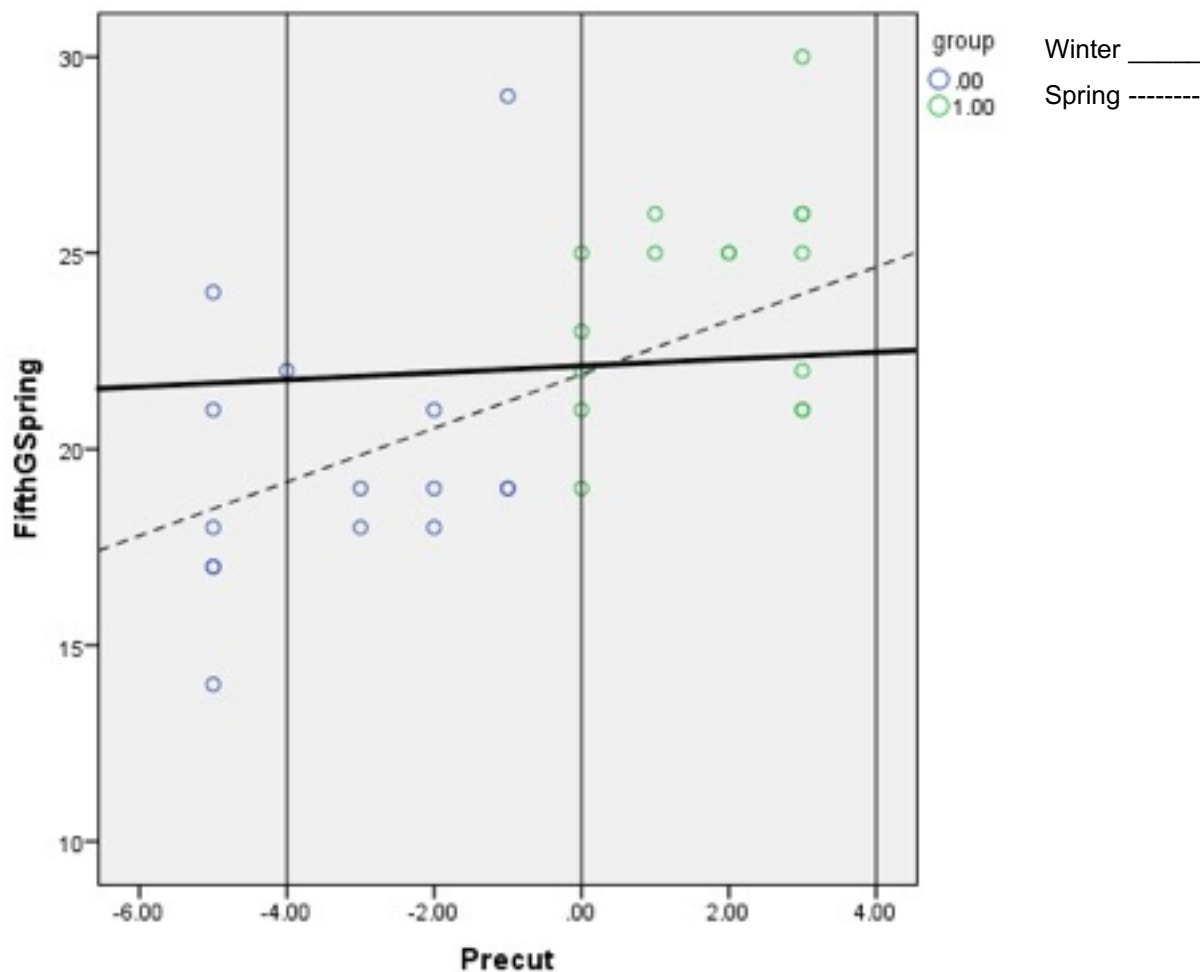


Table 4.10 Effects on Post-Test Scores within Caliper per Grade Level

<u>Group</u>	<u>B</u>	<u>SE B</u>	<u>β</u>	<u>t</u>	<u>p</u>
Third Grade					
Intervention	0.457	0.420	0.172	1.088	0.283
Control	0.887	0.641	0.257	1.383	0.178
Fourth Grade					
Intervention	0.780	1.571	0.198	0.496	0.638
Control	2.896	0.463	0.942	6.253	0.000
Fifth Grade					
Intervention	0.685	0.550	0.327	1.245	0.235
Control	0.706	0.510	0.347	1.384	0.188

* $p < 0.05$, two tailed test

Observer agreement. To control for the standardization of instruction, lessons were observed by the researcher and one other person at the building; the other person was either a Title I teacher or a reading teacher/interventionist. The researcher observed at least one time per week in each classroom and the other observed at least one time per week. The goal was that the researcher and observer should agree on at least 90% of the items on an observation checklist (Appendix E) for modeling (M), guided practice (GP), gradual release of responsibility (GRR), scaffolding (S), and use of non-fiction (NF) text. During the intervention implementation, some of the observers were involved in other end-of-the-year testing and were not available for observation during the intervention time.

Table 4.11 Intervention Lesson Observations--Observer Agreement

Grade/ Classroom	Observer 1						Observer 2						Agreement
	<u>N^a</u>	<u>M</u>	<u>GP</u>	<u>GRR</u>	<u>S</u>	<u>NF</u>	<u>N^a</u>	<u>M</u>	<u>GP</u>	<u>GRR</u>	<u>S</u>	<u>NF</u>	
3-A	3	100%	100%	100%	100%	100%	0	N ^b	N ^b	N ^b	N ^b	N ^b	N/A
3-B	2	100%	100%	100%	100%	100%	3	100%	100%	100%	100%	100%	100%
3-C	3	100%	100%	100%	75%	100%	3	100%	100%	100%	100%	100%	97.5%
3-D	3	100%	100%	100%	100%	100%	3	100%	100%	100%	100%	100%	100%
3-E	3	100%	100%	100%	100%	100%	3	100%	100%	100%	100%	100%	100%
4-A	2	100%	100%	100%	100%	100%	0	N ^b	N ^b	N ^b	N ^b	N ^b	N/A
4-B	2	100%	100%	100%	100%	100%	0	N ^b	N ^b	N ^b	N ^b	N ^b	N/A
5-A	2	50%	100%	100%	100%	100%	1	N ^b	100%	N ^b	N ^b	100%	N/A
5-B	3	100%	100%	100%	100%	100%	2	100%	100%	100%	100%	100%	100%
5-C	2	100%	100%	100%	100%	100%	1	100%	100%	100%	100%	100%	100%

Note: The agreement is based on combined score observations over the duration of the study for each classroom; the percentage is the average of all observations. ^a Number of observations. ^b N-no, not observed. Observer 2 was not able to observe/finish the observation during the lessons for this classroom.

At the conclusion of the study, the teachers participating in the intervention, the Title I teachers/interventionists who were also observers, and administrators were asked to complete a post study survey, the results from which are summarized in Table 4.12. All the respondents indicated that they are planning on using RT in some form of setting. Most viewed RT as a small group or intervention instruction when addressing comprehension.

Table 4.12 Teacher, Observer, and Administrator Exit Survey

Teacher Grade Level	<u>Third</u>	<u>Fourth</u>	<u>Fifth</u>	<u>Observer</u>	<u>Admin</u>		
	4	2	3	3	0		
Years of Teaching	<u>0-5</u>	<u>6-10</u>	<u>11-15</u>	<u>16-20</u>	<u>21-25</u>	<u>25-30</u>	<u>30+</u>
	2	3	1	1	3	1	1
Future Implementation of RT	<u>Whole Group</u>	<u>Small Group</u>	<u>Intervention</u>	<u>Maybe</u>	<u>No</u>	<u>N/A</u>	
	4	8	6	1	0	2	

Summary

Several statistical procedures were used in the analysis of the results, which served as the interpretation and conclusion of findings in the next chapter. The RD in the pilot study received some mis-assignment of scores leading to a “fuzzy” design; the full-scale study design was “sharp”—only the scores of one test were included. For both studies, groups were created in SPSS and split into scores above and scores below a cut-off, then compared in the model by using scores for the second benchmark test. A pre/posttest immediately before and after the intervention was used only in the full-scale study to gauge if the changes were due to the intervention. A one-way ANOVA was used to compare subjects between groups for all students in the pilot and for all students and for students within the caliper in the full-scale. The results were statistically significant for the students in the pilot study and for all fifth graders in the full-scale study. The results were not statistically significant for all third and fourth graders, and all grade levels with scores within the caliper. Although the scores increased, the regression model

detected a significant main effect in the pilot but not in the full-scale study when RT was used in intervention instruction for comprehension for students with similar observable characteristics.

Chapter 5 - Conclusion

Discussion of Findings

The effectiveness of Reciprocal Teaching (RT) as an intervention approach for increasing comprehension was measured statistically during a pilot and a full-scale study. A pilot study was necessary to preliminarily examine the effects of RT. In this case, a small sample was sufficient to evaluate the design, to ensure the validity measures, and to identify potential threats. Based on the results, the effects of the intervention (RT) on the post-test scores were statistically significant for the group that received the intervention but not statistically significant for the control group, or those who did not receive the intervention. In other words, for every unit of the independent variable—scores below the cut off, the dependent variable—post scores on AIMSweb Maze, increased by 1.089 points for the experiment group. Because students were not tested immediately before and immediately after the intervention (within 24 hours) and the “fuzzy” RD design (based on AIMSweb Maze and R-CBM scores), it was unknown whether variables, other than the intervention, may have influenced the scores. This prompted a further investigation, a full-scale study, with a larger sample to determine a more statistically sound explanation for the results of this pilot study.

Besides establishing a “sharp” RD design, the full-scale study implemented a “caliper” or a margin around the cut-off line when comparing the experiment with the control group. For the RD design, it is assumed that students along the cut-off point are similar in academic performance. The opposite is assumed for students at the extreme ends of the regression line above and below the cut-off point—the student who scored the highest is perhaps consistently performing in that range compared to the student who is at the lowest point on the regression line and consistently performing below level. Therefore, a caliper was set four points below and

above the cutoff point at the 25th percentile cut-off, to isolate the scores for students who are assumed to be similar in academic performance.

The full-scale study included a pre-posttest immediately before and after the intervention, as previously suggested by Jacob's and Paris (1985). In contrast to the pilot study, the only variable used for the experiment participants selection was scores below the 25th percentile on AIMSweb Maize only, to ensure a "sharp" RD design. After all consent forms were signed and returned, the full-scale study included 59 students in the experimental groups who worked with classroom teachers during Tier II intervention time.

Effects of RT as Targeted Instruction on Standardized Literacy Assessments

RT had a positive effect on the performance of intermediate grade level students, post intervention, on standardized comprehension tests. However, the overall results suggest that RT was not statistically significant for third and fourth graders but it was statistically significant for fifth graders in the experimental group. Rosenshine and Meister's (1994) meta-analysis also found no significance when RT was used in instruction with third graders.

One possible explanations for the overall results is that third and fourth graders do not have the level of expertise in planning as their older peers, which appears to be the best predictor of efficiency (Brown et, al., 1977; Brown et al.,1983). They could also be considered less capable others within their group level peer interactions (Vygotsky, 1978). The students at these grade levels are also not as receptive to acquiring too many strategies that are introduced in a relatively short duration of instruction, which was found by Baumann (1986) in a study with DRTA and TA. Third grade intervention groups only lasted for three weeks, instead of four, due to prior grade level testing scheduling. In addition, there was a conflict in the last week with more end-of-the-year testing scheduled. There could still be some decoding and fluency issues

that are detected through Oral Reading Fluency (ORF) and RCB-M but these scores were not used to meet the criteria for a sharp RD design for the participants.

Differences in the Effect RT has on Standardized Test Scores at Each Grade

The results indicated that there were differences in effects for each grade level. As recommended in the pilot study, the results were examined further by setting a caliper around the cutoff of the 25th percentile, four points below and above, to compare students whose academic performance is statistically similar. The results of the study when comparing these similar groups of students, indicate that, relative to research questions one and three, the alternative hypothesis that RT has a significant, positive effect on the performance of intermediate grade level students, post intervention, on standardized comprehension tests cannot be accepted; there was a fail to reject the null. This means that for every unit of the independent variable—scores below the cut off, the dependent variable—post scores on AIMSweb Maze, increased by 0.457 points for third, 0.780 points for fourth, and 0.685 points for fifth grade. There were differences in the size of the effect RT has on students' standardized test scores at each grade level, specifically the students around the cut-off margin of four points below and above, which applies to research question two. RT had a positive effect on the performance of intermediate grade level students whose scores were close to the cutoff by a margin of ± 4 , post intervention, on standardized comprehension tests. The alternative hypothesis was not accepted and there was a fail to reject for the null.

There is a difference between statistical significance and practical significance. It is of interest to note the movement of the regression lines at each grade level. The point at which they intercept, in practical terms, is the point at which the effect of the intervention diminishes; up to that point there is a discontinuity between the two lines, suggesting varying positive effects

representing the number of units on posttest scores that increased from the pre-test. As students approach the cutoff, the intervention is still effective but the size of that effect is not as significant because students are closer to the desired score for successful performance and to the benchmark. They do not have to move upwards as far, so even though they make gains, they are not as big as the ones who are further away from the benchmark.

Increase in Overall AIMSweb Maze Scores due to RT as an Intervention

Overall increase in AIMSweb Maze scores was evident when winter benchmark scores were compared to spring benchmark scores. For students who participated in the experimental group, a pre-posttest analysis revealed an interesting pattern when compared to the benchmark scores. Positive gains from the intervention were seen for two students while the rest of the third- grade students had negative gains when comparing pre/post intervention and winter/spring benchmark gain scores. These results for third grade may suggest that the positive gains in student scores may not be due to the intervention. When looking at the post test scores and the spring benchmark score, seven students scored higher on the spring benchmark than they did on the post-test immediately following the intervention and one student's score did not change. However, the rest of the 22 third-graders scored higher immediately following the intervention than they did on the spring benchmark assessment.

For fourth graders, the comparison of post test scores and the spring benchmark scores, one student scored higher on the spring benchmark than on the post-test immediately following the intervention and one student had no change in scores. The rest of the 7 fourth-graders scored higher immediately following the intervention than they did on the spring benchmark assessment. Even with the gains following the intervention, all fourth graders in the experimental

group received scores below the spring benchmark of 19, regardless of whether it was a score immediately after the intervention or at spring benchmark testing.

Fifth grade scores for the experimental group also revealed positive gains from the intervention for four of the twelve students. The rest of the fifth-grade students had negative gains when comparing pre/post intervention and winter/spring benchmark gain scores. For the post test scores and the spring benchmark score comparison, four students scored higher on the spring benchmark than they did on the post-test immediately following the intervention and two students had no change in scores. The rest of the six fifth-graders scored higher immediately following the intervention than they did on the spring benchmark assessment. Three fifth-grade students' post scores immediately following the intervention were at or above the spring benchmark for fifth grade, while only one student remained at or above the spring benchmark at the time of testing.

Importance to the Education Field

The need for effective instruction at just the right time for students who have specific needs in reading is crucial for their future academic, professional, and personal lives. It is of great importance for districts, administrators, and classroom teachers to be informed decision-makers regarding what interventions are implemented that match student needs based on data. Regardless of dwindling school funding, high-stakes testing under new mandates, and technology based instruction, it is still educational professionals at every level that can provide differentiated instruction in the most effective, efficient, and personable way. Knowing if RT is effective, or not as effective, can help educators to make sound instructional decisions, especially for students, for whom time is of the essence for bridging the achievement gap. It may be misleading to assume that the lack of statistical significance in the results indicates that RT is an

ineffective strategy for the Tier II students. Although statistically not significant, the probability of RT ineffectiveness for all third graders is 0.901 or 90% of the time and 0.086 or 8.6% of the time for fourth graders, and 0.05 or 5% of the time for fifth graders in the experimental group. The same analysis for the students who scored within the caliper shows that RT would be ineffective 0.283 or 28.3% of the time, for fourth 0.638 or 63.8% of the time, and for fifth graders 0.235 or 23.5% of the time. It appears that fifth graders would benefit the most from RT as a Tier II intervention, which is promising, considering the new challenges of text complexity and expectations for more independent learning awaiting in middle school.

Limitations

The limitations of the study encompass the size of the sample, consent response, and timing and duration of the study. There were only 59 students who participated in the experiment group; for an RD design, at least 2.5 times the number of a random trial is required in order to reach adequate statistical power. Students who met the criteria but did not return signed consent forms (or permission was not granted), even after teachers followed up in person or via written communication, could not participate in the intervention groups and that also decreased the number of students participating in the experiment group. The lowest number of students who participated was in fourth grade, which could have also affected the non-significant outcomes.

The time of the year also posed several obstacles to the schedule of study duration and observations by a second observer. Being the end of the school year, the study was scheduled around some prior testing obligations and school events. Although several motivational incentives were provided, such as a variety of writing tools, individual folders, and popular animation-characters stickers, students could have been enticed by the spring weather and

anticipation for summer break. Observations in some sites were limited due to staff availability and other testing schedules at that time of the year.

Summarizing can be subjective based on students' background knowledge and perspective on the topic. The interest inventory may have provided the most common interests for the students at each grade level, however, certain topics may not have been appealing to all students in the group. Although it is one of the strategies, students' responses to summarization may have been influenced by lack of interests or by their teacher guidance.

Teacher experience with strategy instruction that is specific to their grade level may have also differed. The intensity and amount of comprehension instruction with younger and older readers could vary. Teachers of transitional readers in third grade still focus on the fluency as much as on the comprehension component of reading, while those in fifth grade work with mostly advanced readers. For these more experienced decoders, the instructional emphasis on comprehension with different types of texts and genre is in-depth and vastly embedded across the curriculum.

Future Research

Possibilities for further research would need to focus on third and fourth grade interventions. Third grade, specifically, is the grade in which students transition from learning to read to reading to learn. Further investigation into the exogenous variable or variables affecting the scores for 3rd graders is necessary, especially for students scoring below the grade level benchmark.

Classroom-based implications. In reference to the RtI framework, interventions should be implemented for all students. The present study evaluated the effects of RT in intervention with students performing below grade-level benchmark and below the 25th

percentile, placing them in need for Tier II intervention. Considerations of implementation should be made for students who are also performing significantly above the grade level benchmarks and need instruction that challenges deeper engagement with content area reading. As with any other initiative, the implementation of RT in Tier II intervention, will require professional development, ongoing support, and continuous conversations in grade level teams and professional learning communities evaluating the success and the progress of student performance.

Teacher training in the use of the strategies with explicit modeling and metacognitive think-aloud are essential for the students to utilize the language and reading behavior patterns during guided practice. Familiarity with gradual release and the ability to realize when scaffolding is appropriate are also two components of the small group instruction that teachers must consider during informal assessments. The goal during small group instruction is to have students take on a more independent role with little teacher support during peer teaching.

The standardized protocol of the lessons and fidelity of implementation during the day and week provide consistency for the intervention. However, with the appropriate teacher training discussed above, the content, text topic and text level, can be differentiated for the students' prior experiences, interests, and maturity. While the parts of the lesson listed on the observation form should be included in the lesson plan, the teacher, following training, is the person who can best make the decision about engaging the students in the process of strategic reading with RT.

Research-based implications. Overall, this study adds to the literature about the effectiveness of RT on improving students' comprehension. The use of quasi-experimental designs in educational research, such as the RD design, is not as prevalent as other, experimental

designs. It is a design that provides for the selection of participants in the experiment based on their need. For an inclusive classroom of diverse learners, addressing students' needs is a professional and ethical responsibility. Teacher experience, the duration of intervention, and a larger sample size are all variables that lend opportunities for further exploration with RT as an intervention in a multivariate analysis with this design.

Exploring the effects with RT as a Tier II intervention across primary grades will pose specific challenges in the criteria for selection in the intervention group. Students in primary grades focus on developing decoding skills and must first be able to read connected text with appropriate fluency before adequate comprehension is achieved. Nevertheless, comprehension should still be a goal at those grade levels.

Examining effects of RT as an intervention can also be the basis for further studies involving students at the other spectrum of academic performance. Instead of working with students below a benchmark on a standardized comprehension test, teachers would plan for tier instruction with students who are performing exceedingly above their grade level or with students who are in a gifted program.

Investigating the effects of the continuity of utilizing RT across the curriculum at the elementary grade levels with whole group and small groups of students is also relevant. How students transfer the use of metacognition from reading to other subjects can be helpful in promoting the process of independent thinking. For students who struggle to rationalize through problem solving in math, draw conclusions in science, or develop perspectives in social studies, learning how to think aloud with a peer may be more developmentally appropriate before internalizing the process.

Conclusion

The overall results suggest that RT had a positive but not statistically significant effect for all third and fourth graders in the experimental group. However, it was statistically significant for all fifth graders in the experimental group. Grade level results for RT effectiveness varied. RT had an insignificant positive effect on the performance of intermediate grade level students, post intervention, on standardized comprehension tests.

The results of the study when comparing similar groups of students, four points above and below the cutoff, indicate that RT had a positive effect on the performance of intermediate grade level students, post intervention, on standardized comprehension tests. The effects of RT were not statistically significant for third, fourth, and fifth graders. RT had a positive effect on these students' scores who were close to the cut off by a margin of ± 4 , post intervention, on standardized comprehension tests. In other words, RT had a positive effect for students with similar observable characteristics.

It appears that fifth grade students benefited the most from RT as an intervention resource for students who struggle with comprehension. The students at this grade level are gaining a resourceful tool that can be helpful when they transition to middle school. The learning environment in middle-school is different from their elementary classroom in several ways: presence of departmentalization, volume of content area reading, and expectations for independent learning. The nature and curriculum of different disciplines at the secondary level, will not take into consideration the comprehension abilities of students. Strategies for reading and learning with text may not be the focus of content coursework. Being able to utilize a strategy, even if it is effective some of the time, can lead to the motivation to persevere when faced with challenging content reading.

For students who are in Tier II intervention, targeted instruction with RT may provide the necessary momentum that will propel their scores above the benchmark. Moving into the field of desired scores for the grade level is a step towards reaching the required mastery of standards. Growth in scores also has another, perhaps even more meaningful point—nurturing the reading development of each child who will one day become the adult, the independent thinker, who is perceptive to the fault and empowerment of the infinite world of old and new literacies.

References

- Adams, A., Carnine, D., & Gersten, R. (1982). Instructional strategies for studying content area texts in the intermediate grades. *Reading Research Quarterly*, 18(1), 27-55.
doi:10.2307/747537
- Alvermann, D. E., Gillis, V. R., & Phelps, S. F. (2013). Planning for content literacy. *Content area reading and literacy: Succeeding in today's diverse classrooms* (7th ed.). Boston, MA: Pearson.
- Armbruster, B. B., Anderson, T. H., & Ostertag, J. (1987). Does text structure/ summarization instruction facilitate learning from expository text? *Reading Research Quarterly*, 22(3), 331-346. doi:10.2307/747972
- Baumann, J. F. (1984). The effectiveness of a direct instruction paradigm for teaching main idea comprehension. *Reading Research Quarterly*, 20(1), 93-115. doi:10.2307/747654
- Baumann, J. F. (1986). Teaching third- grade students to comprehend anaphoric relationships: The application of a direct instruction model. *Reading Research Quarterly*, 21(1), 70-90.
doi:10.2307/747961
- Bean, T., & Steenwyk, F. (1984). The effect of three forms of summarization instruction on sixth graders' summary writing and comprehension. *Journal of Literacy Research*, 16(4), 297-306. doi:10.1080/10862968409547523
- Berkowitz, S. J. (1986). Effects of instruction in text organization on sixth- grade students' memory for expository reading. *Reading Research Quarterly*, 21(2), 161-178.
doi:10.2307/747843

- Brown, A. L., Campione, J. C., & Barclay, C. R. (1979). Training self- checking routines for estimating test readiness: Generalization from list learning to prose recall. *Child Development*, 50(2), 501-512. doi:10.2307/1129429
- Brown, A. L., Day, J. D., & Jones, R. S. (1983). The development of plans for summarizing texts. *Child Development*, 54(4), 968-979. doi:10.2307/1129901
- Brown, A. L., Smiley, S. S., Day, J. D., & Lawton, S. C. (1977). Intrusion of a thematic idea in children's comprehension and retention of stories. *Child Development*, 48(4), 1454-1466. doi:10.2307/1128507
- Bunch, M. B. (2011). Testing English language learners under no child left behind. *Language Testing*, 28(3), 323-341. Retrieved from <http://search.proquest.com.er.lib.k-state.edu/docview/889925065?accountid=11789>
- Carnine, D., Kameenui, E., & Woolfson, N. (1982). Training of textual dimensions related to text- based inferences. *Journal of Literacy Research*, 14(3), 335-340. doi:10.1080/10862968209547459
- Cohodes, S. & Goodman, J. (2012). First degree earns: The impact of college quality on college rates (Faculty Research Working Paper Series No. RWP12-033). Cambridge, MA: Harvard Kennedy School
- Curs, B. R. (2015). What can merit-aid buy? The effects of financial aid packages on the enrollment decisions of applicants to a large public university. University of Missouri working paper. Retrieved from http://web.missouri.edu/~cursb/research/Curs_merit-aid_2015.pdf

- Darling-Hammond, L. (2007). Race, inequality and educational accountability: The irony of "no child left behind". *Race, Ethnicity and Education*, 10(3), 245-260. Retrieved from <http://search.proquest.com.er.lib.k-state.edu/docview/62052580?accountid=11789>
- Davey, B. (1983). Think aloud: Modeling the cognitive processes of reading comprehension. *Journal of Reading*, 27(1), 44-47. Retrieved from <http://AIMSweb.jstor.org.er.lib.k-state.edu/stable/40029295>
- Davey, B., & McBride, S. (1986). Effects of question- generation training on reading comprehension. *Journal of Educational Psychology*, 78(4), 256.
- DesJardins, S., McCall, B., Ott, M., & Kim, J. (2010). A quasi-experimental investigation of how the gates millennium scholars program is related to college students' time use and activities. *Educational Evaluation and Policy Analysis*, 32(4), 456-475. Retrieved from <http://AIMSweb.jstor.org/stable/40963173>
- Dewitz, P., & Graves, M. F. (2014). Teaching for transfer in the common core era. *Reading Teacher*, 68(2), 149-158. doi:10.1002/trtr.1290
- Duke, N. K., Pearson, P. D., Strachan, S. L., & Billman, A. K. (2011). Essential elements of fostering and teaching reading comprehension. In S. J. Samuels, & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (4th ed., pp. 51). Newark, DE: International Reading Association.
- Fisher, D., Brozo, A. G., Frey, N., & Ivey, G. (2007). Reciprocal teaching. *50 content area strategies for adolescent literacy* (pp. 89-92). Upper Saddle River, NJ: Pearson Education, Inc.

- Fisher, D., Frey, N., & Lapp, D. (2009). *In a reading state of mind: Brain research, teacher modeling, and comprehension instruction*. Newark, DE: International Reading Association.
- Fisher, D., & Frey, N. (2014). Contingency teaching during close reading. *Reading Teacher*, 68(4), 277-286. doi:10.1002/trtr.1298
- Fisher, D., & Frey, N. (2015). Teacher modeling using complex informational texts. *Reading Teacher*, 69(1), 63-69. doi:10.1002/trtr.1372
- Fisher, D., Frey, N., & Williams, D. (2002). Seven literacy strategies that work. *Educational Leadership*, 60(3), 70-73. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ655420&site=ehost-live>
- Flavell, H. J. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911.
doi:<http://dx.doi.org.er.lib.k-state.edu/10.1037/0003-066X.34.10.906>
- Flippo, R. F. (2004). Strategy systems for reading, studying, and learning. *Texts and tests: Teaching study skills across content areas* (pp. 88-91). Portsmouth, NH: Heinemann.
- Harris, A. J., & Sipay, E. R. (1980). Improving reading comprehension, II. *How to increase reading ability: A guide to developmental and remedial methods* (7th ed., pp. 479-514). New York, NY: Longman.
- Harvey, S., & Goudvis, A. (2007). *Strategies that work: Teaching comprehension for understanding and engagement* (2nd ed.). Portland, ME: Stenhouse Publishers.
- Havelock, E. A. (1986). *The Muse learns to write: Reflections on orality and literacy from antiquity to the present*. New Haven, CT: Yale University Press.

- Helfeldt, J. P., & Lalik, R. (1976). Reciprocal student- teacher questioning. *The Reading Teacher*, 30(3), 283-287.
- Johns, J. L. & Lenski, S. D. (2014). *Improving reading: Strategies, resources, and common core connections*, (6th ed.). Dubuque, IA: Kendall Hunt Publishing.
- Jacob, A. J. & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *The Review of Economics and Statistics*, 86(1), 226-244.
doi/10.1162/003465304323023778
- Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, 22(3-4), 255-278.
doi:10.1080/00461520.1987.9653052
- Keene, E. O., & Zimmermann, S. (2007). In 2nd (Ed.), *Mosaic of thought: The power of comprehension strategy instruction*. Portsmouth, NH: Heinemann.
- Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2010). Handbook on impact evaluation: Quantitative methods and practices. Washington, DC: World Bank Group. Retrieved from <https://openknowledge.worldbank.org/handle/10986/2693>
- Klingner, J. K., Vaughn, S., & Schumm, J. S. (1998). Collaborative strategic reading during social studies in heterogeneous fourth- grade classrooms. *Elementary School Journal*, 99(1), 3-22.
- Lesik, S. A. (2008). Evaluating developmental education programs in higher education. (ASHE/Lumina Policy Briefs and Critical Essays No. 4). Ames: Iowa State University, Department of Educational Leadership and Policy Studies.

- Lipson, M. Y. & Wixson, K. K. (2009). Getting started in instruction. *Assessment and instruction of reading and writing difficulties: An interactive approach* (4th ed., pp. 554-556). Boston, MA: Pearson Education, Inc.
- Macedo, D. (1994). *Literacies of power: What Americans are not allowed to know*. Boulder, CO: Westview Press.
- Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development*, 48(3), 986-992. doi:10.2307/1128350
- Mier, M. (1984). ERIC/RCS: Comprehension monitoring in the elementary classroom. *The Reading Teacher*, 37(8), 770-774.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010, 2015/2017). English language arts: Anchor standards. *Common core state standards*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers. Retrieved from <http://www.corestandards.org/ELA-Literacy/CCRA/R/>
- National Institute for Direct Instruction. (2015). History of DI research: Beginnings. Retrieved from <https://www.nifdi.org/research/history-of-di-research/beginnings>
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.
- Nolte, R. Y. & Singer, H. (1985). Active comprehension: Teaching a process of reading comprehension and its effects on reading achievement. *The Reading Teacher*, 39(1), 24-31.

- Oczkus, L. D. (2010). *Reciprocal teaching at work: Powerful strategies and lessons for improving reading comprehension*, (2nd ed.). Newark, DE: International Reading Association.
- Oddo, M., Barnett, D., Hawkins, R. O., & Musti-Rao, S. (2010). Reciprocal peer tutoring and repeated reading: Increasing practicality using student groups. *Psychology in the Schools*, 47(8), 842-858. doi:<http://dx.doi.org.er.lib.k-state.edu/10.1002/pits.20508>
- Omanson, R. C., Beck, I. L., Voss, J. F., & Mckweon, M. G. (1984). The effects of reading lessons on comprehension: A processing description. *Cognition and Instruction*, 1(1), 45-67. doi:10.1207/s1532690xci0101_3
- Ong, W. J. (1982/2002). *Orality and literacy: The technologizing of the word*. New York, NY: Routledge.
- Palincsar, A. S., Brown, A. L., Illinois Univ., Urbana Center for the Study of Reading, & Bolt, Beranek and Newman, Inc., Cambridge, M.A. (1983). *Reciprocal teaching of comprehension-monitoring activities. technical report no. 269*. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED225135&site=ehost-live>
- Palincsar, A., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1(2), 117-175.
- Pattison, R. (1982). *On literacy: The politics of the word from Homer to the age of rock*. New York, NY: Oxford University Press.
- Shinn, M. R. & Shinn, M. M. (2002). *AIMSweb training workbook: Administration and scoring of reading maze for use in general outcome measurement*. Eden Prairie, MN: Edformation, Inc.

- Pearson Education, Inc. (2012). AIMSweb technical manual. Bloomington, MN: NCS Pearson, Inc.
- Pearson, P. D., & Gallagher, M. C. (1983). *The instruction of reading comprehension*. (No. 297).
- Raphael, T. E., & Pearson, P. D. (1985). Increasing students' awareness of sources of information for answering questions. *American Educational Research Journal*, 22(2), 217-235.
- Raphael, T. E., & Wonnacott, C. A. (1985). Heightening fourth- grade students' sensitivity to sources of information for answering comprehension questions. *Reading Research Quarterly*, 20(3), 282-296. doi:10.2307/748019
- Raphael, T., & McKinney, J. (1983). An examination of fifth- and eighth- grade children's question- answering behavior: An instructional study in metacognition. *Journal of Literacy Research*, 15(3), 67-86. doi:10.1080/10862968309547490
- Raygor, A. L., & Raygor, R. D. (1985). In Dodd J., Belser J. R. (Eds.), *Effective reading: Improving reading rates and comprehension*. New York, NY: McGraw-Hill, Inc.
- Rinehart, S. D., Stahl, S. A., & Erickson, L. G. (1986). Some effects of summarization training on reading and studying. *Reading Research Quarterly*, 21(4), 422-438. doi:10.2307/747614
- Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research*, 64(4), 479-530.
- Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, 66(2), 181-221.

- Rubin, R. B. (2011). The pell and the poor: A regression-discontinuity analysis of on-time college enrollment. *Research in Higher Education*, 52(7), 675-692.
doi.org/10.1007/s11162-011-9215-6
- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs* (report from the Governing Board of the American Educational Research Association Grants Program). Washington, DC: American Educational Research Association.
- Smith, F. (1988). *Understanding reading* (4th ed., pp. 222-225). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Tovani, C. (2000). Fix it! *I read it, but I don't get it: Comprehension strategies for adolescent readers* (pp. 52-53). Portland, ME: Stenhouse Publishers.
- United States. National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform: a report to the nation and the secretary of education, United States Department of Education. Washington, D.C.: The Commission. Retrieved from http://datacenter.spps.org/uploads/sotw_a_nation_at_risk_1983.pdf
- Uttero, D. A. (1988). Activating comprehension through cooperative learning. *The Reading Teacher*, 41(4), 390. Retrieved from <http://search.proquest.com.er.lib.k-state.edu/docview/203264702?accountid=11789>
- Vacca, R. T., & Vacca, J. A. (2008). From struggling readers to striving readers. *Content area reading: Literacy and learning across the curriculum* (9th ed., pp. 47-53). Boston, MA: Pearson.
- Van Keer, H. & Verhaeghe, J. P. (2005). Effects of explicit reading strategies instruction and peer tutoring on second and fifth graders' reading comprehension and self-efficacy

- perceptions. *Journal of Experimental Education*, 73(4), 291. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ726370&site=ehost-live>
- Vygotsky, L. S. (1985). *Thought and language*. (E. Hanfmann & G. Vakar, Trans.). Cambridge, MA: The M. I. T. Press. (Original work published 1962/1965)
- Vygotsky, L. S. (1978). In Cole M. (Ed.), *Mind in society: The development of higher psychological processes*. Cambridge: Cambridge: Harvard University Press.
- What Works Clearinghouse. (2013). *Reciprocal teaching. what works clearinghouse intervention report*. What Works Clearinghouse. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED544469&site=ehost-live>
- Williams, M. (2001). Making connections: A workshop for adolescents who struggle with reading. *Journal of Adolescent & Adult Literacy*, 44(7), 588-602. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=507765359&site=ehost-live>
- Wixson, K. K. (1983). Questions about a text: What you ask about is what children learn. *The Reading Teacher*, 37(3), 287-293.

Appendix A - Sample AIMSweb Testing Passage

Reading Maze–Test Administration and Scoring

This workbook section covers administration and scoring of Maze and what examiners need to do (1) before testing students, (2) while testing students, and (3) after testing students. Practice tests for scoring are included.

Look at the Maze scores of two third-graders, a typically performing student, and a low performing student. What do you observe?

Emma, A Typical 3rd Grader

Once upon a time there was a merchant whose wife died, leaving him with three daughters.

The two older daughters were good-looking (~~but~~, stand, ~~then~~) very disagreeable. They cared only for (~~until~~, themselves, ~~himself~~) and for their appearance; they spent (~~palace~~, wicked, ~~most~~) of the time admiring their reflections (~~in~~, of, ~~turned~~) a looking glass.

The third and youngest (~~once~~, daughter, ~~gate~~) was quite different from the other (~~him~~, two, ~~beast~~). She was beautiful—so beautiful that (~~I~~, loved, ~~she~~) was known as Beauty. She was (~~also~~, ago, ~~dream~~) good and kind. Everyone loved Beauty, (~~changed~~, by, ~~except~~) for her sisters, who were jealous (~~handsome~~, of, ~~from~~) her. They hated her.

One (~~foot~~, he, ~~day~~) the merchant heard that a ship was (~~expected~~, and, ~~shivered~~) in a far away port with valuable, high, whisper cargo for him. He prepared at (~~she~~, still, ~~once~~) to set off on the long (~~into~~, journey, ~~fountain~~) to the port, for he hoped (~~struck~~, for, ~~to~~) make a great deal of money (~~by~~, of, ~~voice~~) selling the cargo there.

1. What differences did you observe in Emma's and Abby's Maze performance?
2. What other conclusions can you draw?

Abby, A Low-Performing 3rd Grader

Once upon a time there was a merchant whose wife died, leaving him with three daughters.

The two older daughters were good-looking (~~but~~, stand, ~~then~~) very disagreeable. They cared only for (~~until~~, themselves, ~~himself~~) and for their appearance; they spent (~~palace~~, wicked, ~~most~~) of the time admiring their reflections (~~in~~, of, ~~turned~~) a looking glass.

The third and youngest (~~once~~, daughter, ~~gate~~) was quite different from the other (~~him~~, two, ~~beast~~). She was beautiful—so beautiful that (~~I~~, loved, ~~she~~) was known as Beauty. She was (~~also~~, ago, ~~dream~~) good and kind. Everyone loved Beauty, (~~changed~~, by, ~~except~~) for her sisters, who were jealous (~~handsome~~, of, ~~from~~) her. They hated her.

Appendix B - Introductory Lesson Plan

Day 1-Introduction of Strategies

-

Goal (objective): Getting to know four strategies that are used by good readers and help with understanding, or comprehending, what we read--predict, question, clarify, and summarize.

Materials:

*Lexile appropriate expository text, preferably of interests for the students, for read-aloud to introduce each strategy.

*Anchor charts/posters for each strategy

*Group notes folder containing observation chart and prompts chart

Procedure:

1. Anticipatory Set:

Have students turn-and-talk with a partner to discuss what good readers do, then share with the group. (The teacher may also share what they do as a "good"/"grown-up" reader).

2. Objective

State the objective to the group: **“Today, we are going to find out more about four strategies that are used by good readers and help with understanding, or comprehending, what we read. The strategies I will model for you today are to predict, question, clarify, and summarize. In the next few lessons you will get to practice and use these strategies as you work on your reading skills.”**

3. Modeling:
(Some input--show each posters to the students): **“The pictures on each poster are something that reminds us of what each strategy does for us as readers.**

Predictions are sort of like predicting the future; this reminds me of a movie I saw one time about a fortuneteller who predicted the future using a crystal ball. The hand gesture we will use for prediction is (show rounding motion with hands) that will remind us of the "crystal ball."

Questions are something good readers ask all the time. The picture for questions is...question marks of course! That is what we use to show that we are asking a question at the end of a sentence when we write. The gesture for questions is to pretend to hold a microphone (show) as if we are interviewing someone; just like the news people do on TV.

Clarify means to find something we don't know and figure it out or make sense of it . The picture for clarify is a magnifying glass or a spyglass. It will remind us to look for those words, sentences, or parts in the text that don't make sense to us and find a way to figure them out. The gesture for clarify is the sign used by sports players--time out sign (show).

And finally, we will **summarize**. When we summarize, we only pick out the most important parts. When I think of summarizing, I think of when I tell someone about a great movie I have seen--I only tell them about the most important parts of the movie as I only have a few minutes to tell them about a two hour movie! The picture we have for summarize shows how we take a lot of words and sentences and text and only use a few of the most important parts we found in what we read; we only have a small note pad not a whole notebook with a lot of pages.

Notice that each poster will give us "hints" on how to start our sentences when we are **thinking about what we read.**” Keep posters as anchor charts and refer to them as you model; encourage students to refer to these and get comfortable using the specific language for each strategy in the discussions and dialogues.

Explicitly model each strategy, one at a time, by thinking aloud (refer to posters and folder table on sentence starters for each one):

Example for prediction:

Before reading tell students: **I am going to show you how to make predictions and tell you what and how I think about making a specific prediction.**

Think-aloud about the title of the article; quickly skim and scan to look for words that stand out or words that don't seem familiar to you. Point out the picture and the caption under it; tell students that captions can be helpful in helping us to better understand what we are reading about.

Use sentence starters, such as, "**I think I will learn...because....**" Invite students to help you make predictions, ask questions, help clarify, and summarize after each modeling. They are only adding to your thinking--the responsibility of each strategy is still on the teacher. Refer them to the poster for prediction starters.

Example for question:

During reading, read aloud the first paragraph of the article and tell the students that you will either ask an "**I wonder...**" question or a quiz question like the ones that are asked by teachers or a TV game show host. Invite students to ask a shoulder partner a question immediately following your question. Refer them to the poster for question starters. It is acceptable for students to repeat your question.

Example for clarify:

During reading, find a word or a sentence to model on how to clarify. Say, "**I didn't get the [word, sentence, part], so I [reread, read on, chopped the word into smaller parts, thought of what would make sense, etc.].**" Again, refer students to poster for clarify to find other examples for this strategy. If they cannot find examples, have them pretend that they are helping a younger student who has trouble understanding a word, sentence, or a part in the text.

Example for summarize:

Reread and skim/scan the text (or paragraph) and give a quick summary. "**So far...**" or "**This is mostly about....**" Then have students try to summarize; refer to poster for summary but accept student responses that are similar or the same as yours.

Show students **how to predict again** based on what you have read so far, your preview of the next paragraph/section, or your background knowledge on the topic. You can use, "**I think I will learn...because....**" Specifically for non-fiction, model how to look ahead, notice headings, illustrations/photographs with captions, graphs, tables, and how to skim the words in order to make predictions.

4. Guided Practice/Gradual Release:
Continue to do most of the "work" or modeling but invite students to participate after each strategy by sharing with a partner then share with the group. Their participation is what will increase as each paragraph/article is read. Refer them to the posters for sentence starters.

After your think-aloud or during the lesson, have students reflect on the strategies and their use, "**How did I use each reciprocal teaching strategy to help me read this text, and which strategy did you think was the most helpful?**"

5. Informal assessment to guide instruction for the next lesson-record on observation sheet in the group folder as you refer to the rubric.

Pay close attention to how students are applying the use of each strategy:

--Are they using background knowledge or text to predict?

--Are they asking questions that can be answered based on the information in the text or inferred?

--Do they give examples of clarifying words and ideas and ways to clarify?

--Do they summarize the text and include important points/content specific vocabulary?

Appendix C - Basic Daily Lesson Format

Day _____-Lesson: _____

Goal (objective): Using the four strategies with expository text that help with understanding, or comprehending, what we read--predict, question, clarify, and summarize.

Materials:

*Lexile level expository text for each student and for the teacher, preferably of interests for the students:

*Anchor charts/posters for each strategy

*Basic Comprehension Chart-p. 153 in folder and Fab Four Starters p. 268

*For coaching tips, refer to Table 11-in folder and pp. 225-226 in folder

*Group notes folder containing observation chart and prompts chart; Fab Four Placemat, Spinner Game, and Role Tags

*optional--big chart paper, different colors of markers and dry-erase markers, sticky notes

*optional--Spinner or Fab Four Mat

Procedure:

1. Anticipatory Set:

2. Objective:

State the objective to the group: **"Today, we are going to read _____ and use the four strategies to help with understanding, or comprehending, of the text--predict, question, clarify, and summarize."**

Before Reading: Predict and Summarize

During Reading: Question and Clarify

After Reading: Question, Clarify, Summarize, and Check Predictions

(follows the generic lesson plan for small groups--folder/p. 154)

1. Modeling: Discuss the title, text features (headings, tables, charts, captions, font, etc.) to model making initial predictions. Have them make predictions and record on the Basic Comprehension Chart-p. 153, "What We Know" section (can be filled in with dry-erase markers or on the big chart paper).

Ask students what they know about the topic and fill in "Our Predictions."

If in the middle of the text, model summarizing before previewing the next section of the text. Record in chart, "Our Summary."

Model how to use sticky notes to write questions you want answered.

Model finding a part that may be tricky or confusing, mark a C on a sticky and place it by that part. Then model what you use to make sense of it (reread, read on, chopped the word into smaller parts, thought of what would make sense, etc.).

4. Guided Practice/Gradual Release

Encourage students to work in pairs, with your support (scaffold for pairs or individual students) to preview features after you have modeled. Have them find words or parts that would give good ideas for predictions (may have to review how you skim and scan for "key words" or phrases). Have them try making predictions verbally. Ask, **"What do you think you will learn?", "What clues did you use?" or "How do you know?"**

Invite them to ask questions they want answered; scaffold with **"I wonder...."** Add to the chart.

Have students write questions on their sticky notes they would like to have answered or questions others might have. Record on the chart.

Have students work in pairs, with your support (scaffold for pairs or individual students) to find something that seems confusing, mark it with a sticky C, and then refer to the poster or Fab Four Strategy Starters (p. 268) to explain how they made sense or clarified what they did not understand. Record on the chart.

If students are ready to take on more responsibility with these two strategies and the procedures, have them read silently (no round robin) then write questions or clarifications on a sticky and share with their partner and group.

Have students reflect on the strategies and their use, **"How did I use each reciprocal teaching strategy to help me read this text, and which strategy did you think was the most helpful?"**

5. Informal assessment to guide instruction for the next lesson-record on observation sheet. This will guide the next lesson.

How to use the Fab Four Mats: Students can use their dry erase markers to add their predictions, questions, clarify, and summarize. Each student has a different color.

How to use the Spinner: Using a pencil and a large paper clip, have student take turns spinning. Students work in pairs or individually (in later lessons or if ready based on assessment) to

provide either a prediction, question, clarify, or summarize, depending on where the paper clip lands after spinning.

Adapted from *Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension* (second edition) by Lori D. Oczkus. © 2010. Newark, DE: International Reading Association.

Appendix D - Lesson Planning Guide

Reciprocal Teaching Guided Reading Lesson Plan Guide

Fiction Lesson Plan	Nonfiction Lesson Plan
<p>Before Reading Activate prior knowledge</p> <ul style="list-style-type: none"> • "Have you ever...?" • "What do you know about...?" <p>Predict</p> <ul style="list-style-type: none"> • Preview illustrations/chapter books • Skim and scan • "I think...will happen because..." (start graphic organizer) • Setting • Characters • Problem • Events • Ending • Character feelings 	<p>Before Reading Activate prior knowledge</p> <ul style="list-style-type: none"> • "What do you think know or think about...?" <p>Predict</p> <ul style="list-style-type: none"> • Preview illustrations/chapter books • Skim and scan • Predict the following: • What are you wondering about • Use table of contents • Start a graphic organizer, such as a K-W-L chart or Venn diagram. • Write predictions • "I think I will learn...because..."
<p>During Reading Read [silently, chorally, with partners] to find out predictions</p> <p>Question/Clarify</p> <ul style="list-style-type: none"> • Reread to find a point/word to clarify • Or, reread to ask a question • Teacher circulates and models and asks individuals to predict, summarize, question, or clarify 	<p>During Reading Read [silently, chorally, with partners] to find out predictions</p> <p>Question/Clarify</p> <ul style="list-style-type: none"> • Reread to find a point/word to clarify • Or, reread to ask a question • Teacher circulates and models and asks individuals to predict, summarize, question, or clarify
<p>After Reading Predict</p> <ul style="list-style-type: none"> • Return to predictions to confirm or change <p>Question</p> <ul style="list-style-type: none"> • Take turns modeling, then partners ask questions about the characters, events, theme <p>Clarify</p> <ul style="list-style-type: none"> • Discuss/model words or points to clarify and say "I didn't get the part where..., so I..." • Discuss ways to clarify <p>Summarize</p> <ul style="list-style-type: none"> • Summarize as a group or page by page • Fill in a graphic organizer with setting, characters, problem, events, ending, feelings • Which strategy helped you the most today and why? • Give examples 	<p>After Reading Predict</p> <ul style="list-style-type: none"> • Return to predictions to confirm or change <p>Question</p> <ul style="list-style-type: none"> • Model for students as you question using headings/main ideas • Partners ask questions <p>Clarify</p> <ul style="list-style-type: none"> • Discuss/model words or points to clarify and say "I didn't get the part where..., so I..." • Discuss ways to clarify <p>Summarize</p> <ul style="list-style-type: none"> • Summarize as a group, or page by page with partners • Return to fill in graphic organizer with what was learned • Which strategy helped you the most today and why? • Give examples

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Oczkus.
 ©2010, Newark, DE: International Reading Association. May be copied for classroom use.

Appendix E - Observation Form

Observation Check Sheet

Date:	Observer:			
Classroom: A B (circle)				
Modeling				
Guided Practice				
Gradual Release				
Scaffolding				
Feedback from provided chart				
Nonfiction Text				
Anchor Charts Prompts visible/referred to				
Strategy	Predict	Question	Clarify	Summarize
Teacher				
Student 1				
Student 2				
Student 3				
Student 4				
Student 5				

Notes:

Appendix F - Interest Inventory

News About Me

A News Story About _____
(write your name here)

News About My Family

I have _____ brothers and sisters.
They are _____ years old.
I like to play with _____.
My mother and I like to _____.
My father and I like to _____.
I (like/do not like) to play alone.
I help at home by _____.
The thing I like to do at home is _____.

News About My Pets

I have a pet _____.
I (do/do not) take care of my pet.
I do not have a pet because _____.
I would like to have a pet _____.

News About My Books and My Reading

I like to read about _____.
The best book I ever read was _____.
I (do get/do not get) books from the library.
I have _____ books of my own at home.
I read aloud to _____.
My _____ reads to me.

News About My Friends

My best friend is _____.
I like (him/her) because _____.
We play _____.
I would rather play (at my house/at my friend's house) because _____.

News About Things I Like and Dislike

I do not like _____.
I like _____.
I am afraid of _____.
I am not afraid of _____.

News About My Wishes

When I grow up, I want to be _____.
If I could have three wishes I would wish
(1) _____
(2) _____
(3) _____

page 1

From Jerry L. Johns and Susan Davis Lenski, *Improving Reading: Strategies and Resources* (3rd ed.). Copyright © 2001 Kendall/Hunt Publishing Company (1-800-247-3458, ext. 5). May be reproduced for noncommercial educational purposes.

News About My Travels and Adventures

I have traveled by:

_____ bus _____ car
_____ airplane _____ truck
_____ boat _____ train
_____ bicycle _____ van

I have visited these interesting places:

_____ circus _____ zoo
_____ farm _____ park
_____ hotel _____ museum
_____ bakery _____ library
_____ airport _____ fire station
_____ factory, and _____

The best adventure I ever had was _____

News About My School Subjects

My favorite subject is _____.

The subject I dislike most is _____.

I am best at _____.

I wish I was better in _____.

Write any other news about yourself below.

page 2

News About My Hobbies and Collections

One of my best hobbies is _____

I collect _____.

I want to collect _____.

My other hobbies are _____

Movie, Radio, and Television Favorites

I see _____ movies each week.

I like to listen to _____ on the radio.

I see _____ television programs a day.

My favorite programs are _____

Appendix G - Teacher Feedback Chart

Predicting

When predicting with fiction, students

- Preview the front and back covers, illustrations, and headings before reading
- Predict what is likely to happen next based on clues from the text or illustrations
- Use what they know (from text and prior knowledge)
- Stop to predict during reading
- Continue to make logical predictions based on clues from the text

When predicting with nonfiction, students

- Preview the front and back covers, illustrations, and headings before reading
- Predict what is likely to be learned based on clues from the text or illustrations
- Apply what they already know to help make a prediction
- Stop to predict during reading
- Continue to make logical predictions based on clues from the text

When using metacognition with either fiction or nonfiction, students tell how predicting helps them understand the text.

The language of prediction that students use^a may include the following phrases:

- I think...because....
- I imagine...because....
- I'll bet...because....
- I suppose...because....
- I wonder if...because....
- I predict....because....

Questioning

When questioning with fiction, students

- Ask questions based on the text (that is, the answers are in the text)
- Ask questions that are based on the main idea or question of the story
- Ask some detail-oriented questions
- Ask some inferential questions

When questioning with nonfiction, students

- Ask questions based on the text (that is, the answers are in the text)
- Ask questions that are based on the main idea of the reading
- Ask some detail-oriented questions
- Ask questions based on nonfiction text features, such as maps, captions, and diagrams
- Ask some inferential questions

When using metacognition for either fiction or nonfiction, students can tell how asking questions helps them to understand the text.

The language of questioning that students use may include the following words:

- Who
- Why
- What
- How
- When
- What if
- Where

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Oczkus.
© 2010. Newark, DE: International Reading Association. ^aSee Mowery, S. (1995). *Reading and writing comprehension strategies*. Harrisburg, PA: Instructional Support Teams. May be copied for classroom use.

Clarifying

When clarifying with fiction, students

- Express confusion with specific portions of text, such as ideas or events, that are difficult to understand
- Identify words that are difficult to pronounce or understand

When clarifying with nonfiction, students

- Point out confusing ideas related to the content of the reading
- Point out confusing portions of text, such as sentences, paragraphs, and pages
- Identify words that are difficult to pronounce or understand

When using metacognition for either fiction or nonfiction, students

- Give strategies for clarifying words
- Tell strategies for clarifying ideas
- Tell how clarifying helps them understand text

The language of clarifying that students use may include the following phrases:

- I didn't understand the part about..., so I [see list below]
- This doesn't make sense, so I [see list below]
- I can't figure out..., so I [see list below]

So I...

- | | |
|---|--|
| • Reread | • Checked the parts of the word I knew |
| • Read on for clues | • Blended the sounds of the word |
| • Reread the sentence to see if it made sense | • Tried another word |

Summarizing

When summarizing fiction, students

- Retell the story in their own words and include the setting, characters, problem, key events, and resolution

Or, they

- Give only key points in a short one- or two-sentence summary
- Summarize in a logical order
- Reread to remember main ideas
- Refer to illustrations to retell or summarize the text

When summarizing nonfiction, students

- Retell the key points or ideas
- Leave out unnecessary details
- Summarize in a logical order
- Reread to remember main ideas
- Refer to illustrations, headings, and other text features to retell or summarize the text

When using metacognition for either fiction or nonfiction, students tell how summarizing helps them understand the text.

The language of summarizing that students use may include the following words or phrases:

- | | | |
|----------------|--|---------------------------|
| • First,.... | • The most important ideas in this text are:.... | • A key event is when.... |
| • Next,.... | • The story takes place.... | • This part is about.... |
| • Then,.... | • The main characters are.... | • This book is about.... |
| • Finally,.... | • A problem occurs when.... | |

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Oczkus.
© 2010. Newark, DE: International Reading Association. May be copied for classroom use.

Table 11
Guided Reading Coaching Strategy, Protocol, and Prompts

Selecting a strategy to coach

Move around the group and choose strategies to coach based on any of the following criteria:

- Coach every student on the focus strategy you just modeled in a minilesson.
- Coach each student on a different strategy—the one he or she needs to practice based on your observations. Use a note-taking chart.
- Ask each student which strategy he or she wants to practice!
- Coach the strategy that naturally fits the portion of text the student is reading (e.g., a logical place to make a prediction).

Coaching protocol

1. Student reads a small portion of text aloud to you (e.g., sentence, paragraph, page).
2. Select a strategy to coach (see above choices).
3. Model for the student if necessary (e.g., "Watch me...").
4. Guide using language prompts for the strategy (e.g., "I didn't get...so I...").
5. Ask the student to tell you what he or she just learned (e.g., "I learned to summarize using the headings.").

Predict



I think...will happen next, because....
I think I will learn....
I bet....
I predict....
Maybe....

Tell me what you think will happen next in the text. Why do you think that is so?
How is this text organized? Based on that, what do you think will happen next?
Look at the headings to make a prediction.
Study the pictures to make your prediction.
Skim the words in the text. What do you think this is about?

Question



I wonder...
Who, what, when, where, why, or how.
What if...
Why do you think...?

What is a question that you could ask about this page? What is the answer? How did you get that answer?
Ask a wonder question starting with "I wonder...?"
Ask a quiz question: who, what, when, where, why, or how.
Ask a thinking question that starts with "Why do you think...?"

Clarify



I didn't get the [word, sentence, part, chapter] so I...[reread, read on, looked for parts I know, broke into syllables, substituted a synonym]

Identify a difficult word, sentence, or part. How did you figure it out? Give at least two ways.
What is a word or idea that would be difficult for a younger child to read?
Show me a place where you had to read twice to get a picture in your head.
Is there a metaphor or figure of speech that you need to clarify?

Summarize



This is mostly about....
So far....
First..., Next..., Then..., Finally....
In the [beginning, middle, end],....
The theme is....

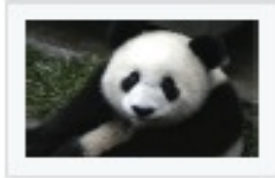
Summarize this [paragraph, page, chapter]. Or, summarize what we have read so far.
What is this mainly about?
Is there a problem to be solved? If so, what?
Tell me the main idea. Use the topic sentence or heading to help you.

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Oczkus. © 2010. Newark, DE: International Reading Association. May be copied for classroom use.

Appendix H - Non-Fiction Text Sample

Meet a Baby Panda

By ReadWorks



In the United States, pandas live in zoos. Some baby pandas are born in the zoo.

At first, a baby panda does not look like its mother. It does not have much hair at all. Its eyes are closed. The baby drinks its mother's milk. The mother holds the baby in her paws.

The baby grows quickly. Now it looks like its mother. Soon the little panda starts to eat bamboo. Bamboo is a kind of plant.

Then the panda spends less time with its mother. It likes to be on its own. It will climb trees. It will sit on their branches.

Bread Around the World



Tortillas cooking

Bread has been an important food for people for a very long time. The first bread was made thousands of years ago. People back then made dough with crushed up grain and water. Then they baked the dough on hot stones. This made a flat kind of bread.

For a long time, flat breads were the only kind of bread. Then, people in Egypt learned how to make bread rise. They also made ovens for baking the bread. Ever since then, there have been many kinds of bread. Some are flat. Some are puffy. All of them are eaten and enjoyed in different cultures! Here are some different breads that are eaten by people around the world:

A tortilla is a popular kind of bread in Mexico. It is flat and round. This bread can be made from cornmeal or wheat flour.


Another kind of round, flat bread is a chapati. This bread is often eaten in India and other countries in South Asia. It is also made from wheat flour.

Challah bread was first made and eaten by Jewish people. It is very different from tortillas and chapatis. It is not flat. Instead, it's puffy. Also, the dough is braided before it is baked. So the bread has a very special shape!

What other kinds of breads do you know about?

Appendix I - Anchor Charts

Predict




I think....
 I'll bet....
 I wonder if....
 I imagine....
 I suppose....
 I predict....
 I think I will learn...because....
 I think...will happen because....

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Ocklick
© 2010 Newark, DE: International Reading Association. May be copied for classroom use.

275

Question

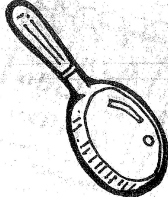


I wonder...?
 Who...?
 What...?
 When...?
 Where...?
 Why...?
 How...?
 What if...?
 Why do you think...?
 How do you think...?

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Ocklick
© 2010 Newark, DE: International Reading Association. May be copied for classroom use.

276

Clarify



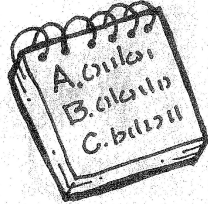
I didn't get the [word, idea, part] so I [reread, read on, sounded it out, etc.].
 I didn't understand the part where....
 This doesn't make sense.
 This is a tricky word because....
 To clarify, I can

- Reread
- Read on
- Sound out
- Look for parts I know
- Blend the sounds together
- Think of another word
- Make a picture in my head
- Talk to a friend

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Ocklick
© 2010 Newark, DE: International Reading Association. May be copied for classroom use.

277

Summarize




The most important ideas are....
 This part was mostly about....
 This [book, chapter, part, article] was about....
 [First, Next, Then, Finally],
 The story takes place....
 The main characters are....
 A problem occurs when....
 In the [beginning, middle, end],

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Ocklick
© 2010 Newark, DE: International Reading Association. May be copied for classroom use.

278

Appendix J - Role Cards


211




PREDICTOR

PREDICTOR
(predicts what the story or text will be about)

1. Ask your group to look at the cover, next heading, or next illustration.
2. Encourage group members to discuss what they see.
3. Ask all group members to write their predictions on their Literature Discussion Sheet for Reciprocal Teaching.
4. Share your predictions first. Then, encourage all group members to share their predictions with reasons for them.




Reprinted by permission of Linda Ward Beech, *Teaching to Think: Powerful Strategies and Lessons for Improving Reading Comprehension* (second edition) by Linda D. Cook, © 2010, New York, DE: International Reading Association. May be copied for classroom use.



QUESTIONER

QUESTIONER
(asks questions about the text—what, where, when, why, how, and what if?)


1. Ask your group members to write one or two questions that could be answered by reading this text.
2. Ask your questions first. Call on a volunteer to answer your question.
3. Ask for other volunteers to ask their questions. Don't allow yes or no questions!



Reprinted by permission of Linda Ward Beech, *Teaching to Think: Powerful Strategies and Lessons for Improving Reading Comprehension* (second edition) by Linda D. Cook, © 2010, New York, DE: International Reading Association. May be copied for classroom use.

212


213




CLARIFIER

CLARIFIER
(finds areas where a word or idea needs to be explained—reread, read aloud, use what you know, break a word into chunks, and think about what makes sense)

1. Ask the group to read this portion of text and look for confusing ideas or words. (What if you had to explain the book to a kindergarten?)
2. Ask the group to write one confusing or difficult word or idea.
3. Share your word or idea first. Tell how you figured it out using the Clarifying Bookmarks.
4. Ask for volunteers to give their words and ideas. Ask how they figured them out. If someone has a difficult word or idea that he or she didn't figure out, ask group members for ways to clarify the unclear word or idea.




Reprinted by permission of Linda Ward Beech, *Teaching to Think: Powerful Strategies and Lessons for Improving Reading Comprehension* (second edition) by Linda D. Cook, © 2010, New York, DE: International Reading Association. May be copied for classroom use.



SUMMARIZER

SUMMARIZER
(tells in own words a brief summary of the important point or points)

1. Ask the group to summarize this part of the text in writing on the Literature Discussion Sheet for Reciprocal Teaching.
2. Share your summary first.
3. Ask if anyone would like to add to your summary or give his or her own.

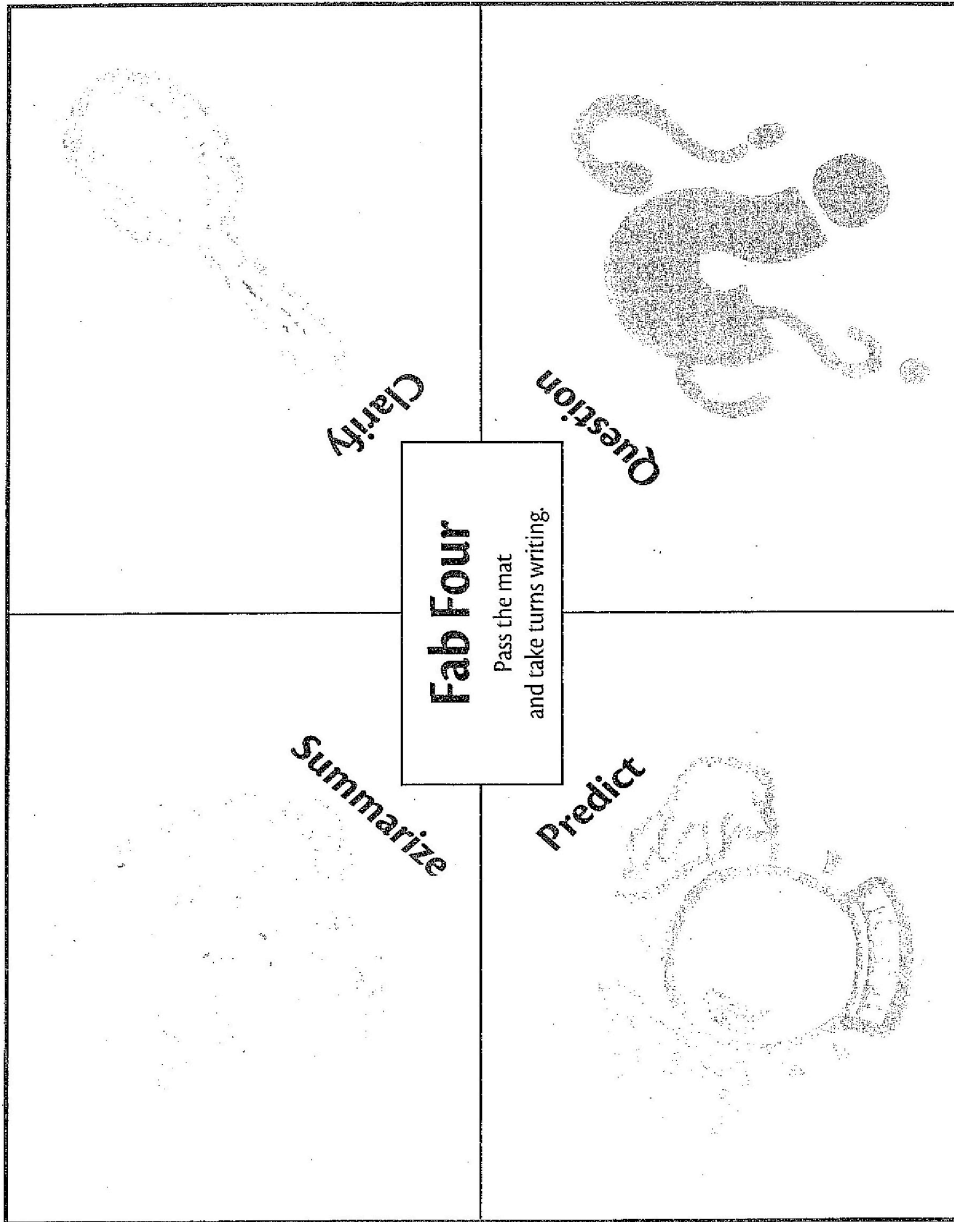


Reprinted by permission of Linda Ward Beech, *Teaching to Think: Powerful Strategies and Lessons for Improving Reading Comprehension* (second edition) by Linda D. Cook, © 2010, New York, DE: International Reading Association. May be copied for classroom use.

214

Appendix K - RT Dry-Erase Mat

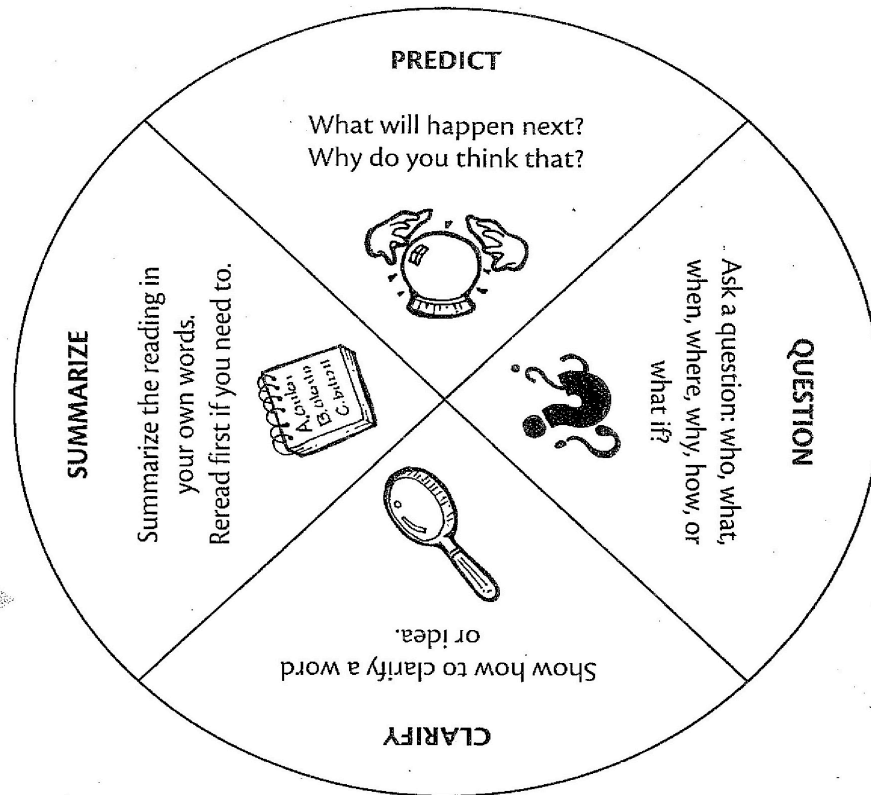
Fab Four Mat



Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Ozkhus © 2010, Newark, DE: International Reading Association. This graphic adapted from Bennett, B. & Rolheiser, C. (2001). *Beyond Monet: The Artful Science of Instructional Integration*. Toronto, ON: Barrie Bennett. May be copied for classroom use.

Appendix L - RT Spinner Game

Reciprocal Teaching Spinner



When you land on a strategy that someone else already landed on, you can...	
PREDICT <ul style="list-style-type: none"> • Give a new prediction. • Or, add a detail to the last one. 	CLARIFY <ul style="list-style-type: none"> • Choose another word or idea to clarify. • Or, give more ways to clarify the same word or idea.
QUESTION <ul style="list-style-type: none"> • Ask another question that begins with a different question word. 	SUMMARIZE <ul style="list-style-type: none"> • Give a new summary. • Or, add to the previous summary.

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Oczkus. © 2010, Newark, DE: International Reading Association. May be copied for classroom use.

Appendix M - Student/Teacher Graphic Organizer

Basic Comprehension Chart for Guided Reading Groups

What We Know (our experiences, background knowledge, and/or connections with the text)	
Our Predictions (make predictions before reading and check their accuracy after reading)	Our Questions (who, when, what, where, why, how, and what if)
Our Words or Ideas to Clarify	How We Clarified
Our Summary	

Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension (second edition) by Lori D. Oczkus.
© 2010. Newark, DE: International Reading Association. May be copied for classroom use.